材料物理基础

第2版

任凤章　主编

机械工业出版社
CHINA MACHINE PRESS

VISIT…

LANZAROTE Caliente.com

材料物理基础

第2版

任凤章　主编

机 械 工 业 出 版 社

材料物理是介于物理学与材料学之间的一门边缘学科，所涉及面极广，它旨在利用物理学中的一些基础知识来阐明材料的各种物理性能和转变过程。本书突出基础，略去一些过于深奥的专业阐述，突出物理学的主干，从物理学的一些基本概念、基本原理和基本定律出发，阐述材料本身的结构、性质和它们在一些外界条件下发生的变化及其变化规律。本书共分 9 章。前 3 章扼要介绍固体材料物理所涉及的量子力学概念、统计力学基本知识、经典物理学方面的波动现象、原子结构、化学键、晶体结构及金属中的电子态，为没有学过固体物理的读者提供一些基础知识。其第 4 ~ 第 7 章主要介绍材料的热、电、磁及介电性能的物理本质。第 8 章介绍第一性原理的基础理论、计算方法和应用。第 9 章介绍原子间相互作用对势、多体势和晶格反演势。

本书可作为材料科学与工程类专业硕士研究生的教材或参考书，也可供材料科学与工程领域的大专院校教师和科技工作者参考。

图书在版编目（CIP）数据

材料物理基础/任凤章主编. —2 版. —北京：机械工业出版社，2012. 5
ISBN 978-7-111-39868-4

Ⅰ. ①材…　Ⅱ. ①任…　Ⅲ. ①材料科学—物理学—高等学校—教材
Ⅳ. ①TB303

中国版本图书馆 CIP 数据核字（2012）第 226365 号

机械工业出版社（北京市百万庄大街 22 号　邮政编码 100037）
策划编辑：张秀恩　　责任编辑：张秀恩
版式设计：霍永明　　责任校对：刘志文
封面设计：陈　沛　　责任印制：乔　宇
三河市宏达印刷有限公司印刷
2012 年 5 月第 2 版 · 第 1 次印刷
169mm×239mm · 21. 75 印张 · 423 千字
标准书号：ISBN 978-7-111-39868-4
定价：39.00元

凡购本书，如有缺页、倒页、脱页，由本社发行部调换

电话服务	网络服务
社服务中心：（010）88361066	教材网：http://www.cmpedu.com
销售一部：（010）68326294	机工官网：http://www.cmpbook.com
销售二部：（010）88379649	机工官博：http://weibo.com/cmp1952
读者购书热线：（010）88379203	**封面无防伪标均为盗版**

第 2 版前言

本书第 1 版 2006 年出版，2008 年第二次印刷，被国内多所高校选为研究生教材。教学实践表明，本书的总体结构和内容深浅程度是合适的，但也存在着一些问题，主要是未对目前广泛应用的一些新理论进行介绍。因此，根据材料物理的发展，对本书进行修订再版。

第一性原理计算，是基于量子力学理论在电子水平上对材料进行计算设计。因其不含任何可调（经验或拟合）参数，就可以通过求解薛定谔方程，得到材料的电子结构，预测材料的组分、结构与性能，设计具有特定性能的新材料，从而在材料科学研究中具有重要的地位。原子间相互作用势是所有有关原子水平上的计算机模拟的基础，广泛应用于物理、化学、材料科学中，在材料的缺陷、表面和界面、断裂、团簇的结构及其他温度特性的计算模拟中起着重要的不可替代的作用。因此，本书在原有内容的基础上增加了第一性原理计算和原子间相互作用势两部分内容，以适应材料学科快速发展的需要。

本书这次修订主要在两个方面：第一，对第 1 版章节中个别的难以理解的句子重新书写，并对一些不妥之处进行修改。第二，增加了两章新内容，分别是第 8 章的第一性原理计算和第 9 章的原子间的作用势。

新增内容第 8 章和第 9 章由河南科技大学龙永强编写。第 1 章～第 5 章修订由河南科技大学任凤章完成，第 6 章、第 7 章修订由河南科技大学龙永强完成。河南科技大学文九巴教授对本书的修订给予了很大的支持和帮助。

虽然本书进行了修订，但不妥之处仍难免，敬请广大读者继续提出宝贵意见。

河南科技大学　任凤章

2011 年 8 月

第 1 版前言

材料物理是介于物理学和材料学之间的一门边缘学科，它旨在从物理的微观角度来讨论固体材料的各种性能。要了解材料的性能，不能只满足于实验研究，还必须深入研究材料的内部组织、结构。这样，就要涉及到组成材料的原子运动、电子态、杂质和缺陷及相结构等问题，将材料的各种性能提到微观理论的高度来研究，为改善材料的各种性能和发展新材料奠定基础。因此，本书突出物理学的主干，从物理学的一些基本概念和基本原理出发，解释材料在一些外界条件下出现的一些物理现象和效应。本书可作为材料科学与工程类专业硕士研究生的教材或参考书，也可供材料科学与工程领域的大专院校教师和科技工作者参考。

材料涉及的范围很广，其中的固体材料就包括金属、高分子、无机非金属材料，每种材料的物理性能又是多种多样。因而，材料物理涉及的范围也很广，本书遴选固体材料物理中一些最基础、最基本的内容。全书共分 7 章，各章内容简介分述如下：第 1 章概括性地叙述了固体材料物理所涉及的量子力学概念、统计力学基本知识及经典物理学方面的波动现象。为非物理专业的读者接受微观现象中的量子效应奠定基础。第 2 章介绍材料中原子结构、化学健和晶体结构的基本理论。第 3 章介绍金属的电子理论，重点讨论量子自由电子理论和能带理论。第 4 章首先介绍晶格振动的经典理论和量子理论，在此基础上，重点讨论晶体的热容、热膨胀和热传导的机理及变化规律。第 5 章介绍金属、半导体和超导体电导的物理基础。第 6 章介绍电介质在不同频率、不同场强的电场作用下所表现出的现象，物质内部的电极化过程以及铁电体的非线性介电、压电、热电效应等。第 7 章介绍了原子磁矩、物质磁性的分类、铁磁性的分子场理论、亚铁磁性的超交换理论，也介绍铁磁性物质内部的能量和磁畴的形成。因材料科学与工程类专业的一些教科书和参考书对晶体中的扩散和相变及晶体缺陷都有较详细的叙述，本书略去本该列入的此部分内容。

本书由河南科技大学任凤章博士主持编写，各章节分别由河南科技大学贾淑果博士（第 1 章）、任凤章博士（第 3 章，第 4 章，第 5 章，第 6 章），清华大学李锋军博士（第 2 章 2.1、2.1 节，第 7 章），河南科技大学祝要民博士（第 2 章 2.3 节）撰写。河南科技大学刘平教授、田保红教授和文九巴教授对本书的撰写给予了高度关注，并提出了许多好的建议。全书由任凤章统稿。编写过程中参考了一些相关的文献，已在书后列出，谨此对有关学者表示感谢，如有疏漏，敬请

包涵。

感谢河南科技大学学科建设办公室和材料科学与工程学院对本书出版的大力支持。本书获河南科技大学学术著作出版基金资助。

由于编者水平所限，加之时间仓促，书中的缺点、错误恐怕不少，恳请读者批评指正。

编　者

2006 年 8 月

目　　录

第 2 版前言

第 1 版前言

第 1 章　固体物理基础……1

1.1　量子力学引论……1

1.1.1　早期量子论……1

1.1.2　德布罗意物质波……12

1.1.3　测量和测不准原理……16

1.1.4　薛定谔方程……19

1.2　波动现象……22

1.2.1　波的基本方程式……22

1.2.2　谐波……23

1.2.3　变数分离……24

1.2.4　波的叠加……26

1.2.5　傅里叶解析……29

1.2.6　波束（波包）……31

1.2.7　波束的运动……32

1.3　统计力学概要……33

1.3.1　系统的微观运动状态描述……33

1.3.2　分布和微观状态……36

1.3.3　玻尔兹曼分布……38

1.3.4　玻色分布和费米分布……40

第 2 章　原子结构、化学键及晶体结构……42

2.1　原子结构……42

2.1.1　氢原子……42

2.1.2　核外电子的排布规律及周期表中各元素原子的电子层结构……51

2.2　分子结构与化学键……56

2.2.1　原子间作用力及由来……57

2.2.2　化学键……60

2.3　晶体结构……68

2.3.1　晶体学基础……68

2.3.2　典型晶体结构……75

第 3 章　金属电子论……82
3.1　经典自由电子论……82
3.1.1　经典自由电子论的基本假设……82
3.1.2　经典自由电子论的主要成就……82
3.1.3　经典自由电子论遇到的困难及原因……84
3.2　量子自由电子理论……84
3.2.1　阱内势能为零的电子状态……85
3.2.2　三维自由电子……87
3.2.3　量子自由电子理论对金属导电及电子摩尔热容的解释及遇到的问题……91
3.3　周期势场和布洛赫（A. Bloch）定律……92
3.3.1　周期势场……92
3.3.2　布洛赫定律……93
3.3.3　克勒尼希—彭尼（Kronig-Penney）模型……95
3.4　能带理论……98
3.4.1　准自由电子近似……98
3.4.2　紧束缚电子近似（原子轨道线性组合法）……105
3.5　能带中的电子运动……110
3.5.1　电子速度 $v(k)$与波矢 k 的关系……110
3.5.2　有效质量 m^*与波矢大小 k 的关系……112
3.6　能带理论的简单应用……113
3.6.1　导体、半导体和绝缘体的能带……113
3.6.2　金属的能带……115
3.6.3　合金中的能带……118
第 4 章　晶格振动和晶体的热学性质……121
4.1　晶格振动的经典理论……121
4.1.1　一维简单晶格振动……122
4.1.2　一维复式格子振动……124
4.1.3　声学波和光学波……126
4.1.4　周期性边界条件（玻恩—冯卡门（Born-von-Karman）边界条件）·128
4.2　晶格振动的量子化……129
4.2.1　一维简单晶格振动的总能量……129
4.2.2　三维复式格子的总能量……131
4.3　固体的热容……132
4.3.1　晶态固体热容的经验定律和经典理论……133
4.3.2　晶态固体热容的量子理论……134

4.3.3　固体热容与组织及状态的关系……138
4.4　固体的热膨胀……140
4.4.1　晶格振动的非谐效应——热膨胀的物理本质……141
4.4.2　热膨胀系数与其他物理量之间的关系……142
4.4.3　固体材料热膨胀的一些规律……144
4.5　固体热膨胀的反常现象……146
4.5.1　因瓦效应的微观机制……147
4.5.2　因瓦合金的一些特性……149
4.6　固体的热传导……149
4.6.1　固体材料热传导的宏观规律……149
4.6.2　固体传热的微观机理……150
4.6.3　影响热导率的因素……155
第 5 章　材料的电导性能……160
5.1　金属材料的电导性能……160
5.1.1　金属材料电导机制与马基申定理……160
5.1.2　影响纯金属导电的因素……161
5.1.3　固溶体合金的导电特性……165
5.1.4　金属化合物、中间相及多相合金导电性……167
5.2　半导体导电性能……169
5.2.1　本征半导体……169
5.2.2　杂质半导体……172
5.2.3　半导体的表面能级……177
5.2.4　半导体接触……178
5.3　超导体……185
5.3.1　超导态三个重要特征……186
5.3.2　超导态三个性能指标……188
5.3.3　两类超导体……189
5.3.4　超导现象的物理本质……189
第 6 章　材料的介电性能……192
6.1　电介质的极化……193
6.1.1　极化现象及物理量……193
6.1.2　宏观极化强度与微观极化率的关系（在静电场中）……194
6.1.3　电介质极化机制……197
6.2　交变电场下的电介质行为……207
6.2.1　复介电常数和介质损耗……207

6.2.2 电介质弛豫和频率响应 ……208
6.2.3 介电损耗分析 ……211
6.3 电介质在电场中的破坏 ……213
6.3.1 热击穿机制 ……213
6.3.2 电击穿机制 ……215
6.3.3 局部放电击穿 ……217
6.4 铁电性 ……217
6.4.1 铁电体、电畴 ……218
6.4.2 钛酸钡自发极化的微观机理及电畴的形成……219
6.4.3 铁电体的临界性质 ……221
6.5 压电性和热释电性 ……223
6.5.1 压电性 ……223
6.5.2 热释电性 ……228
第 7 章 材料的磁性能 ……230
7.1 磁学基本量及物质磁性分类 ……230
7.1.1 磁学基本量 ……230
7.1.2 物质的磁性分类 ……231
7.2 原子和离子的固有磁矩 ……232
7.2.1 孤立原子（离子）本征磁矩……232
7.2.2 固体中的原子（离子）磁矩……236
7.3 物质的抗磁性和顺磁性 ……238
7.3.1 抗磁性 ……238
7.3.2 顺磁性 ……238
7.3.3 抗磁性金属与顺磁性金属……240
7.4 铁磁体（包括铁氧体）自发磁化……241
7.4.1 外斯（P. E. Weiss）铁磁性假说 ……241
7.4.2 直接交换作用 ……242
7.4.3 间接交换作用 ……244
7.5 铁磁体中的磁晶各向异性、磁致伸缩……248
7.5.1 磁晶各向异性能 ……248
7.5.2 退磁场能 ……250
7.5.3 磁致伸缩 ……250
7.6 畴壁与磁畴结构 ……251
7.6.1 畴壁 ……252

7.6.2 磁畴 ……253
7.6.3 单畴结构 ……256
第 8 章 第一性原理计算 ……257
8.1 引言 ……257
8.2 第一性原理的基础理论 ……258
8.2.1 多原子体系的薛定谔方程（Schrödinger function） ……258
8.2.2 密度泛函理论（DFT） ……262
8.2.3 交换关联能泛函 ……265
8.2.4 赝势（Pseudopotential） ……267
8.3 第一性原理的计算方法 ……271
8.3.1 波函数的展开 ……271
8.3.2 自洽场（SCF）运算 ……275
8.3.3 结构优化 ……278
8.3.4 能带计算 ……280
8.3.5 晶体物理性质的计算 ……281
8.4 常用的第一性原理计算软件 ……283
8.4.1 CASTEP 软件包 ……283
8.4.2 WIEN 软件包 ……284
8.4.3 ABINIT 软件包 ……284
8.5 第一性原理方法在材料科学研究中的应用 ……285
8.5.1 Cu-Ni-Si 合金析出相稳定性的第一性原理研究 ……285
8.5.2 BN 电子结构的第一性原理计算 ……291
8.5.3 TiB 弹性性质的第一性原理计算 ……294
第 9 章 原子间相互作用势 ……296
9.1 原子间相互作用势概述 ……296
9.2 原子间相互作用对势的形式 ……297
9.2.1 “刚球”模型 ……297
9.2.2 Buckingham 势 ……298
9.2.3 Lennard-Jones（L-J）势 ……299
9.2.4 Morse 势 ……299
9.2.5 Bom-Mayer 势 ……300
9.2.6 多项式形式的势 ……300
9.3 原子间相互作用多体势模型 ……301
9.3.1 Tersoff 和 Brenner 势 ……301
9.3.2 Stillinger-Weber 势 ……304

9.3.3　Slater−Kirkwood 类型的三体势……305
9.3.4　Chelikowsky−Philips−Kamal−Strauss 势……306
9.3.5　Khor−Das Sarma 势……306
9.3.6　Biswas−Hamann 势……307
9.3.7　Murrell−Mottram 势……307
9.3.8　Erkoc 势……308
9.4　嵌入（埋入）原子势（Embedded−atom method，EAM）……308
9.4.1　原子对势理论与准服子理论和有效介质理论……308
9.4.2　原型 EAM 理论……309
9.4.3　分析型嵌入原子（AEAM）理论……311
9.4.4　改进分析型 EAM 模型（MAEAM）……312
9.5　Finnis−Sinclair（F−S）多体势……324
9.5.1　F−S 多体势的初始模型……324
9.5.2　改进的 F−S 多体势模型……325
9.5.3　立方样条函数的 F−S 多体势模型……326
9.5.4　合金中的 F−S 多体势……326
9.6　晶格反演方法（LIM）……329
9.6.1　Carlsson−Gelatt−Ehrenreich（CGE）公式……329
9.6.2　LIM 与异种原子间的对势反演……330
9.6.3　LIM 与原子间的多体势反演……331
9.6.4　陈氏三维晶格反演（Chen−Möbius）……332
参考文献……335

第 1 章　固体物理基础

固体物理所采用的方法，是从原子论的观点出发，把观察到的各种宏观现象（导电性、介电性、磁性等）看成是构成一切物质的原子或原子集团的行为，应用量子力学和统计力学对它们作统一的论述。因此，本章将扼要介绍后面固体材料物理所涉及的量子力学概念和统计力学基本知识。同时，本章也将简要介绍经典物理学方面的波动现象，有助于理解量子论中的波动现象。

1.1　量子力学引论

量子力学是与 20 世纪一起来到人间的。这门理论为什么叫力学，“力学”这个名词的物理学的广阔定义是物质的运行原理。其次，为什么叫量子？在拉丁文里，量子的意思是“分立的部分”或“数量”。量子力学所讨论的就是物质世界的诸种特性中的“分立性”。量子力学是解释构成一切物质的原子和原子集团所表现的宏观现象（电导率、介电常数、磁导率等）的行为的科学。

量子力学的建立，揭示了微观世界的基本规律，使人们对自然界的认识产生了一个飞跃，为原子物理学、固体物理学和粒子物理学的发展奠定了理论基础。同时，量子力学也深入到化学，生物学等其他领域，形成了量子化学、分子生物学等边缘学科。量子力学的诞生，还大大推动了新技术的发展，如晶体管、集成电路、激光、超导材料等，促进了生产力的提高。

19 世纪末，物理学理论已发展到相当完善的阶段。物体的机械运动在速度远比光速小时准确地遵从牛顿力学的规律；电磁现象的规律（包括光的波动理论）总结为麦克斯韦方程组；热现象也有完整的热力学和经典统计物理的理论。当时有不少物理学家认为，物理学的大厦已经基本建成。正当物理学家为经典物理学理论的辉煌成就而欢欣鼓舞之际，一些新的实验事实与经典物理学理论发生尖锐的矛盾。例如，迈克耳孙—莫雷实验没有提供绝对静止的状态存在的证据，黑体辐射的“紫外灾难”等，使经典物理面临新的危机，以致开尔文（L. Kelvin）把这些矛盾说成是在物理学晴朗天空的“两朵乌云”。相对论和量子论的建立驱散了这两朵乌云，开创了物理学中一场深刻的革命。

1.1.1　早期量子论

1900 年，普朗克（M. Planck）为了解释黑体辐射问题，提出了振荡偶极子的

能量是以不连续的量值改变的量子假设，并导出了与实验结果相符合的黑体辐射公式。以后在爱因斯坦（A. Eintesin）、康普顿（A. H. Compton）的理论中进一步提出了“光子学说”，认识到电磁辐射以微粒的形式吸收和发射，从而揭示了光具有微粒和波动的双重特性，称为光的波粒二象性。在普朗克、爱因斯坦明确了光的量子性以后，1913 年玻尔（N. Bohr）根据卢瑟福（E. Rutherford）的原子有核模型以及原子光谱的规律性，提出氢原子的量子论，初步奠定了原子物理学的基础。但是，由于当时对微观粒子的基本属性缺乏认识，玻尔的理论本身仍有其不可克服的缺点。

直到 1924 年，德布罗意（L. de Broglie）在光具有波粒二象性的启发下，认为微观实物粒子（如电子、质子、中子等）与光子一样，也具有波粒二象性，提出了物质波的假说，这一假设不久为电子衍射实验所证实。1925 年至 1932 年间，薛定谔（E. Schrödinger）、海森伯（W. K. Heisenberg）、玻恩（M. Born）、狄拉克（P. A. M. Daric）等人在物质波假设的基础上建立起描写微观实物粒子运动的新理论，称为量子力学。

1. 热辐射和普朗克的能量子假设

（1）绝对黑体热辐射实验

任何固体或液体，在任何温度下都在发射各种波长的电磁波，这种由于物体中的分子、原子受到热激发而发射电磁波的现象称为热辐射。物体向四周所发射的能量称为辐射能。实验表明，热辐射具有连续的辐射能谱，波长自远红外区延伸到紫外区。并且辐射能按波长的分布主要决定于物体的温度。

在单位时间内，从物体表面单位面积上所发射的波长在 λ 到 $\lambda+\mathrm{d}\lambda$ 范围内的辐射能 $\mathrm{d}M_\lambda$，与波长间隔 $\mathrm{d}\lambda$ 成正比，比值 $\mathrm{d}M_\lambda/\mathrm{d}\lambda$ 称为单色辐出度，用 M_λ 表示。单位时间内从物体表面单位面积上所发射的各种波长的总辐射能，称为物体的辐出度，用 $M(T)$表示。在一定温度 T 时，物体的辐出度与单色辐出度的关系为

$$M(T)=\int_0^\infty M_\lambda(T)\mathrm{d}\lambda \tag{1-1}$$

根据基尔霍夫（G. R. Kirchoff）辐射定理，只要知道黑体的辐射度以及物体的吸收比，就能了解一般物体的热辐射性质。因此，从实验和理论上确定黑体的单色辐出度就是研究热辐射问题的中心。绝对黑体就像质点、刚体、理想气体等模型一样，也是一种理想化的模型。可以用不透明材料制成开小孔的空腔，作为在任何温度下能 100%地吸收辐射能的黑体模型。可用实验方法测定绝对黑体的单色辐出度 $M_{0\lambda}(T)$。图 1-1 表示黑体在不同温度 T 下的 $M_{0\lambda}(T)$随 λ 变化的实验曲线。

（2）绝对黑体热辐射理论解释

1）维恩（W. Wien）公式　在经典物理学中，把组成黑体空腔壁的分子或原子看作带电的线性谐振子。1896 年，维恩假设，平衡辐射时黑体辐射的能量按频

率分布，和同温度的理想气体分子按动能的麦克斯韦分布规律类同，从而得出辐射空腔中能量密度分布函数 $M_{0\lambda}(T)$为

$$M_{0\lambda}(T) = c_1\lambda^{-5}\mathrm{e}^{-\frac{c_2}{\lambda T}} \tag{1-2}$$

式中，c_1 和 c_2 为常数；λ 为波长（$v=c/\lambda$，v 为频率，c 为光速）。此式为维恩（W. Wien）公式。维恩（W. Wien）公式在波长较短、温度较低时，才与实验结果相符，而在长波区与实验曲线相差很大，如图 1-2 所示（某一温度下）。

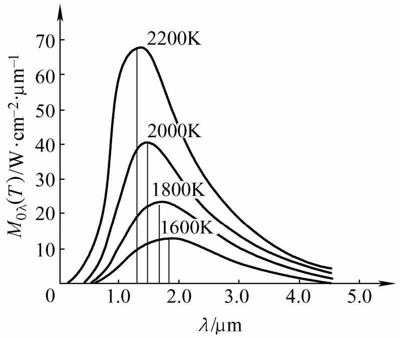


图 1-1　绝对黑体的辐出度按波长分布曲线

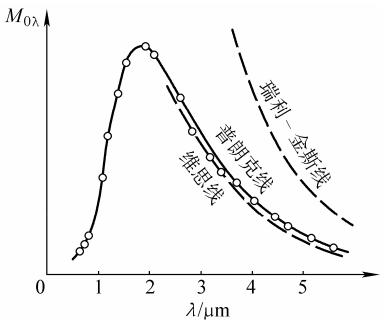


图 1-2　黑体辐射的理论公式与实验结果比较（∘—实验结果）

2）瑞利（Lord Rayleigh）-金斯（J. H. Jeans）公式　维恩的假设是缺乏根据的，其理论结果与实际不符并未引起人们的重视和惊奇。1900 年至 1905 年间，瑞利和金斯认为辐射空腔内的电磁辐射，由于腔壁反射形成很多驻波。根据经典电磁理论，他们算出单位体积内，频率在 v 到 v+dv 之间振动的数目为$(8\pi v/c^3)\mathrm{d}v$，这里 c 为光速。如果认为经典统计方法也适用于电磁辐射，那么，按经典统计能量均分定理，每一个振动方式的平均能量为 kT（k 是玻尔兹曼常数），这样在 v 至 v+dv 之间的能量为

$$M_{0\lambda}(T)\mathrm{d}v = \frac{8\pi v^2}{c^3}kT\mathrm{d}v \tag{1-3}$$

即

$$M_{0\lambda}(T) = \frac{8\pi v^2}{c^3}kT \tag{1-4}$$

式（1-4）称为瑞利-金斯公式。这个公式虽然在低频部分与实验符合，但由于辐射的能量与频率的平方成正比，所以随频率增大而单调增加，在高频部分趋于无限大，即在紫端发散，后来这个失败被埃伦菲斯特（B. Ehrenfest）称为“紫外灾难”。瑞利-金斯公式是用经典物理学的方法来研究热辐射所得的结果，与实验结果不相符合，明显地暴露了经典物理学的缺陷。因此，开尔文（Lord Kelvin）认为黑体

辐射实验是物理学晴朗天空中一朵令人不安的乌云。

3）普朗克公式和能量子　为解决维恩公式和瑞利-金斯公式的困难，普朗克利用内插法将适用于短波的维恩公式和适用于长波的瑞利-金斯公式衔接起来，在 1900 年提出了一个新的公式

$$M_{0\lambda}(T) = 2\pi hc^2\lambda^{-5}\frac{1}{e^{\frac{hc}{k\lambda T}}-1} \tag{1-5}$$

式中，h 后来称为普朗克常数。这一公式称为普朗克公式。它与实验结果符合得很好。由普朗克公式不难得到维恩公式。当波长很短或温度较低时，普朗克公式可近似写成维恩公式；当波长很长或温度很高时，普朗克公式可近似写成瑞利-金斯公式。

普朗克得到上述公式后，他指出“即使这个新的辐射公式证明是绝对精确的，如果仅仅是一个侥幸揣测出来的内插公式，它的价值也只能是有限的。”因此，他要寻找这个公式的理论根据。经过深入研究和分析，他发现必须使谐振子的能量取分立值，才能得到上述普朗克公式。由此他提出以下的假设：把辐射黑体分子、原子的振动看作谐振子，这些谐振子可以发射和吸收辐射能。但是这些谐振子只可能处于某些分立的状态。在这些状态中，谐振子的能量并不像经典物理学所允许的可具有任意值。相应的能量是某一最小能量 ε（ε 称为能量子）的整数倍，即 $n\varepsilon$（n 为正整数，称为量子数）。对于频率为 ν 的谐振子，最小能量为 $\varepsilon=h\nu$。在辐射或吸收能量时，振子从这些状态中的一个状态跃迁到另一个状态，即振子只能“跳跃式”地辐射能量。

从经典物理学可知，原子、分子振动能量遵守玻尔兹曼分布，如果按照普朗克的量子假设，则频率为 ν 的谐振子能量为 $nh\nu$ 的概率正比于 $e^{-nh\nu/kT}$。因而谐振子的平均能量

$$\overline{\varepsilon} = \frac{\sum_{n=0}^{\infty} nh\nu e^{-\frac{nh\nu}{kT}}}{\sum_{n=0}^{\infty} e^{-\frac{nh\nu}{kT}}} = \frac{h\nu}{e^{\frac{h\nu}{kT}}-1} \tag{1-6}$$

又根据经典电动力学理论得到黑体的单色辐出度为

$$M_{0\lambda}(T) = \frac{2\pi\nu^2}{c^2}\overline{\varepsilon} \tag{1-7}$$

将式（1-6）代入式（1-7）即得普朗克公式。

由此可见，正是黑体辐射的实验事实迫使普朗克作出了能量子的假设，这样的假设是与经典物理学的概念格格不入的。从经典物理学看来，能量子的假设是荒诞的、不可思议的。一直到 1905 年，爱因斯坦在普朗克能量子假设的基础上提

出光量子概念，正确地解释了光电效应，从此普朗克能量子假设才冲破经典物理思想的束缚，逐渐为人们所接受。由于普朗克发现了能量子，对建立量子理论作出了卓越贡献，获 1918 年诺贝尔物理学奖。

2. 光电效应和爱因斯坦的光量子假设

（1）光电效应的实验规律

光照射到金属表面时，电子从金属表面逸出的现象称为光电效应。这一效应是赫兹（H. R. Hertz）在 1887 年研究电磁波的波动性质时偶然发现的。他用两套放电电极做实验，一套产生振荡，发出电磁波；另一套充当接收器。电极之间存在火花放电的间隙。此后，汤姆孙（W. Thomson）和勒纳德（P. Lenard）测出光电子的荷质比与阴极射线的荷质比相近，这就肯定光电流和阴极射线实质相同，都是高速运动的电子流。

一个研究光电效应的实验装置如图 1-3 所示。在一抽成高真空度的容器内，装有阴极 K 和阳极 A。阴极 K 为金属板。当单色光通过石英窗口照射到金属板 K 上时，金属板便释放出电子，这种电子称为光电子。如果在 A、K 两端加上电势差 U，则光电子在加速电场作用下，飞向阳极，形成回路中的光电流。实验结果可归纳如下。

1）饱和电流　实验指出：以一定强度的单色光照射电极 K 时，加速电势差 U（$U_A - U_K$）愈大，光电流 I 也愈大。当电势差增加到一定量值时，光电流达饱和值 I_H 如图 1-4 所示。这意味着从电极 K 发射出来的电子全部飞到 A 极上。实验也表明饱和电流 I_H 随入射光的发光强度的增大而增大，也就是说单位时间内，受光照的金属板释放出来的电子数和入射光的发强度成正比。

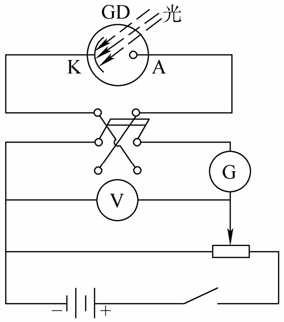


图 1-3　光电效应实验装置图

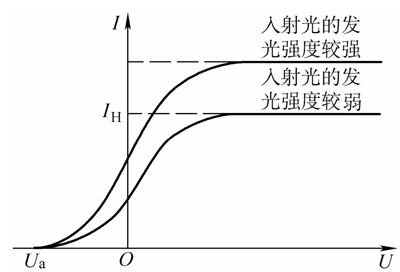


图 1-4　光电效应伏安特性曲线

2）遏止电势差　降低加速电势差的量值，光电流 I 也随之减小。当电势差 U 减小到零并逐渐变负时，光电流 I 并不等于零，这表明从金属板 K 释出的电子具有初动能。如果使负的电势差足够大，从而使由金属板 K 表面释出时具有最大速度 v_m 的电子也不能到达 A 极时，光电流便降为零。光电流为零时，外加电势差的

绝对值 U_a 叫做遏止电势差。光电子逸出时的最大动能和遏止电势差的关系为

$$\frac{1}{2}mv_m^{\ 2}=eU_a \tag{1-8}$$

式中，m 和 e 分别为电子的质量和电荷量。遏止电势差与入射光的光强度无关，也就是说电子的初动能与入射光的光强度无关。

3）遏止频率（又称红限） 实验表明遏止电势差 U_a 与入射光频率 v 有关。它们的关系如图1-5实验曲线所示，不同的曲线对应于不同的阴极金属。这一关系为线性关系，可用数学式表示为

$$U_a=Kv-U_0 \tag{1-9}$$

式中，K 是直线的斜率，是与金属材料无关的一个普适恒量。对不同金属来说，U_0 的量值不同；对任何一金属，U_0 为恒量。将式（1-8）代入式（1-9）得

$$\frac{1}{2}mv_m^{\ 2}=eKv-eU_0 \tag{1-10}$$

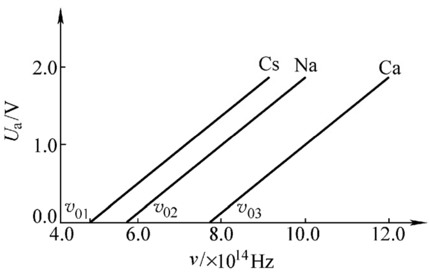


图1-5 截止电压与频率的线性关系

式（1-10）表明光电子的最大初动能随入射光的频率 v 的减少而线性减少。频率降到某一值 v_0 时，U_a 降到零，这时光电子的初动能为零，不再发生光电效应。这个频率 v_0 称为光电效应的红限频率。不同的金属有不同的红限频率。

实验还发现，光电子的发射时间与光的强弱无关，光电子的逸出几乎是在光照到金属表面上即时发生的，其延迟时间在 10^{-9}s 以下。

（2）光的波动说遇到的困难

经典理论认为光是在一定频率范围内的电磁波，当光照到金属表面时，金属中的电子将从入射光中吸收能量，从而逸出金属表面，逸出时的初动能应决定于入射光的发光强度，即决定于光振动的振幅，而与光的频率无关。但实验结果是，任何金属所释放出的光电子的初动能都随入射光的频率线性地增加，而与入射光的发光强度无关。

按照波动说，如果发光强度足够供应从金属释出光电子所需要的能量，那么光电效应对各种频率的光都会发生。但实验事实是，每种金属都存在一个红限频率 v_0。

按照波动说，光的能量是均匀地分布在整个波前上，金属中的电子吸收能量的范围是有限的。金属中的电子从入射光波中吸收能量，必须积累到一定的量值，才能释出电子。显然入射光愈弱，能量积累的时间就愈长，但实验结果并非如此。

由此可见，光的波动理论难以解释光电效应的实验规律。

（3）爱因斯坦的光量子假设

爱因斯坦从普朗克的能量子假设中得到了启发，他认为普朗克的理论只考虑了辐射物体上谐振子能量的量子化，即谐振子所发射或吸收的能量是量子化的，他假定空腔内的辐射能本身也是量子化的，就是说光在空间传播时，也具有粒子性，想象一束光是一束以光速 c 运动的粒子流，这些粒子称为光量子，现称为光子。每一光子的能量也就是 $\varepsilon = h\nu$，不同频率的光子具有不同的能量。

按照光子理论，光电效应可解释如下：当金属中一个自由电子从入射光中吸收一个光子后，就获得能量 $h\nu$，如果 $h\nu$ 大于电子从金属表面逸出时所需的逸出功 A，这个电子就可从金属中逸出。根据能量守恒定律，应有

$$h\nu = \frac{1}{2}m{v_{\mathrm{m}}}^2 + A \tag{1-11}$$

式中，$\frac{1}{2}m{v_{\mathrm{m}}}^2$ 是光电子的最大初动能，式（1-11）称为爱因斯坦光电效应方程。爱因斯坦方程表明光电子的初动能与入射光频率之间的关系。从而解释了式（1-10）。入射光的发光强度增加时，光子数也增多，这就很自然地说明了饱和电流或光电子数与光的发光强度之间的正比关系。假定 $\frac{1}{2}m{v_{\mathrm{m}}}^2 = 0$，由方程式（1-11）得 $\nu_0=A/h$。这表明如果光子频率低于 ν_0（红限），不管光子数目多大，单个光子没有足够的能量去发射光电子。同样由光子理论可以得出，当一个光子被吸收时，全部能量立即被吸收，不需要积累能量的时间，这也就自然地说明了光电效应的瞬时发生的问题。

由于爱因斯坦发展了普朗克的思想，提出了光子假说。成功地说明了光电效应的实验规律，荣获 1921 年诺贝尔物理学奖。

关于光的本性的研究，麦克斯韦和赫兹肯定了光是电磁波，爱因斯坦又提出了光是粒子——光子。综合起来，近代关于光的本性的认识是：光既具有波动性，又具有粒子性，即光具有波粒二象性。光的波动性用光波的波长 λ 和频率 ν 描述，光的粒子性用光子的质量、能量和动量描述。按照量子理论，光子的能量 $\varepsilon = h\nu$。根据相对论的质能关系 $\varepsilon = mc^2$。光子的质量 $m = h\nu/c^2 = h/(\lambda c)$，那么，光子的动量 $p=h/\lambda$。

（4）康普顿散射

1）实验规律　1923 年康普顿（A. H. Compton）研究了 X 射线经物质散射的实验，进一步证实了爱因斯坦的光子概念。图 1-6 是康普顿实验装置的示意图。X 射线源发射一束波长为 λ_0 的 X 射线，并投射到一块石墨上，经石墨散射后，散射束穿过光阑，其波长及相对强度可以由晶体和探测器所组成的摄谱仪来测定，改变散射角，进行同样的测量。康普顿发现，在散射光谱中除有与入射线波长 λ_0 相同的射线外，同时还有波长 $\lambda > \lambda_0$ 的射线。这种改变波长的散射称为康普顿效应。康普顿因发现此效应而获得 1927 年诺贝尔物理学奖。

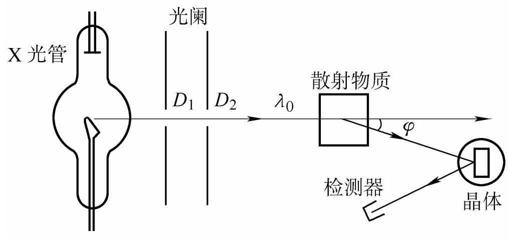


图 1-6　康普顿实验装置

2）光子理论的解释　按经典电磁理论，光的散射是这样产生的：当电磁波通过物体时，将引起物体中带电粒子作受迫振动，从入射波吸收能量，而每个振动着的带电粒子，将向四周辐射电磁波。从波动观点来看，带电粒子受迫振动的频率等于入射光的频率，所发射的光的频率应与入射光的频率相同。可见，光的波动理论能够解释波长不变的散射而不能解释康普顿效应。如果应用光子的概念，并假设单个光子和实物粒子一样，能与电子等粒子发生弹性碰撞，那么康普顿效应能够在理论上得到与实验相符合的解释。根据光子理论，一个光子与散射物中的一个自由电子或束缚微弱的电子发生碰撞后，散射光子将沿某一方向进行，这一方向就是康普顿散射的方向。在碰撞过程中，一个自由电子吸收一个入射光子能量后，发射一个散射光子，当光子向某一方向散射时，电子受到反冲而获得一定的动量和能量。在整个碰撞过程中，动量守恒和能量守恒。因此，散射光子的能量就比入射光子的能量低。因为光子的能量与频率之间有关系 $\varepsilon = h\nu$，所以散射光的频率要比入射光的频率小。如果光子与原子中束缚得很紧的电子碰撞，光子将与整个原子作弹性碰撞，因原子的质量要比光子大很多，按照碰撞理论，散射光子的能量不会显著地减小，因而散射光的频率也不会显著地改变。康普顿偏移非常小，所以观察到散射线里也有与入射线波长相同的射线。

X 射线的散射现象，在理论上和实验上的符合，不仅有力地证实了光子理论，说明了光子具有一定的质量、能量和动量，而且这个现象所研究的，不是整个光束与散射物体的作用，而只是个别光子与个别电子间的作用，所以这种现象同时也证实了能量守恒和动量守恒两定律，在微观粒子相互作用的基元过程中，也同样严格地遵守着。

3. 氢原子光谱和玻尔的电子角动量量子化假设

（1）氢原子光谱的实验规律

1884 年巴耳末（J. J. Balmer）研究了埃格斯特朗（A. J. Ångström）从氢气放电管中获得的氢原子的 4 条明亮的光谱线，这些谱线分别称为 H_α（λ=6562.10Å，1Å=0.1nm）、H_β（λ=4860.74Å）、H_γ（λ=4340.10Å）、H_δ（λ=4101.20Å）。巴耳末发现这四条光谱线的波长可以用一个简单的数学公式表示出来，即

$$\lambda = B\frac{n^2}{n^2-4} \quad (n=3, 4, 5, \cdots) \tag{1-12}$$

式中，B=3645.6Å。由式（1-12）计算所得的波长数值在实验误差范围内与测得的数值是一致的，后人称这个公式为巴耳末公式，而将它所表达的一组谱线称为巴耳末系。

1890 年，里德伯（R. J. R. Rydberg）将巴耳末公式改变为用波长的倒数来表示的形式。如果令 $\tilde{\nu} = 1/\lambda$，$\tilde{\nu}$ 称为波数。巴耳末公式可改写为

$$\tilde{\nu} = \frac{4}{B}\left(\frac{1}{2^2}-\frac{1}{n^2}\right) = R_{\mathrm{H}}\left(\frac{1}{2^2}-\frac{1}{n^2}\right) \quad (n=3, 4, 5, \cdots) \tag{1-13}$$

式中，R_{H}=4/B，称为里德伯常数。后来氢原子光谱其他谱线系也先后被发现，人们将巴耳末公式表达为更广泛的形式

$$\tilde{\nu} = R_{\mathrm{H}}\left(\frac{1}{m^2}-\frac{1}{n^2}\right) \tag{1-14}$$

式中，m=1, 2, 3, …；对每一个 m，n=m+1，m+2，m+3，…构成一个谱线系。

1920 年伍德（W. Wood）拍摄了巴耳末系的 22 条谱线并进行了测量，证明了巴耳末公式的正确性。后来发现氢原子光谱的其他谱线系也与巴耳末公式符合得较好。在氢谱线实验规律的基础上，里德伯、里兹（W. Ritz）等人在研究其他元素（如一价碱金属）的光谱时，发现碱金属光谱也可分为若干线系，其频率或波数也有和氢谱线类似的规律性。

原子光谱线系可用这样简单的公式来表示，且其结果又非常准确，这说明它深刻地反映了原子内在的规律。

（2）卢瑟福原子模型及其与经典电动力学理论的矛盾

1）卢瑟福原子模型　1909 年，卢瑟福（E. Rutherford）、盖革（H. Geiger）和马斯顿（E. Marsden）用镭作放射源，进行 α 粒子穿射金属箔的实验，发现入射束中多数粒子仍保持其原来的方向，但也有不少粒子偏转了很大角度。卢瑟福意识到这种大角散射不可能是由于很多小偏离的累积效应。1910 年底，卢瑟福在 α 粒子散射实验的基础上提出了核式结构模型。他认为原子中有一个体积很小的核，其直径约为 10^{-12}～10^{-13}cm，约为原子直径的 $1/10^4$～$1/10^5$，但却几乎集中了原子的全部质量，带负电的轻得多的电子则在很大的空间里绕核运动。而且一定元素的原子核上的正电荷数目等于核外电子数。为了检验这一模型是否符合观察到的散射结果，必须根据电学定律和力学定律导出一个公式能算出 α 粒子在离排斥中心不同距离处通过时偏转角的大小。这个公式指出偏离原来运动方向角 θ 的 α 粒子数应当与 $\sin(\theta/2)$的 4 次方成反比。这个结论与观察到的散射曲线非常相符，这就证明了原子的核型模型的正确性。

2）卢瑟福原子模型与经典电动力学理论的矛盾　卢瑟福核型模型一开始就由

于同经典电动力学理论矛盾尖锐而遇到了困难。因为按照经典理论，电子绕核运动是加速的，会自动向外辐射电磁波，从而丧失能量而逐渐落向原子核。在这一过程中，电子绕核转动频率连续改变，应向外发射连续光谱。这些推论同原子的稳定性以及原子发射线状光谱的事实相矛盾。于是人们就遇到一个十分棘手的问题，如何说明实在原子的稳定性？这个稳定性问题就成为 20 世纪摆在物理学家面前的几个困难问题之一。

（3）玻尔的电子角动量量子化假设

卢瑟福原子模型被实验证实是正确的，玻尔接受了此模型。为了解决模型同经典电动力学理论矛盾尖锐而遇到的困难，1913 年，玻尔在卢瑟福的核型结构的基础上，把量子化概念应用到原子系统，提出三个基本假设作为他的原子理论的出发点。

1）定态假设　原子系统中的电子只能处在一系列不连续的能量状态，在这些状态中，虽然电子绕核作加速运动，但并不辐射电磁波，这些状态称为原子系统的稳定状态（简称定态）。

2）频率条件　当电子从一个能量为 E_n 的定态轨道跃迁到另一个能量为 E_k 的定态轨道时，就要发射或吸收一个频率为 v_{kn} 的光子。光子的频率与两定态能量的关系为

$$v_{kn} = \frac{|E_n - E_k|}{h}$$

式中，h 为普朗克常数。当 $E_n > E_k$ 时发射光子，$E_n < E_k$ 时吸收光子。

3）角动量量子化条件　在电子绕核作圆周运动中，其稳定状态必须满足电子的角动量 L 等于 $\frac{h}{2\pi}$ 的整数倍的条件，即

$$L = mvr = n\hbar = n\frac{h}{2\pi} \tag{1-15}$$

式中，n 为正整数，称为量子数；m 为电子质量；v 为电子绕核速率；r 为电子绕核轨道半径；$\hbar$（$=\frac{h}{2\pi}$）称为约化普朗克常数。

玻尔根据上述假设很好地解释了氢光谱规律。

（4）氢原子的能级和光谱

由玻尔的这些假设很容易求得氢原子在稳定态中的能量。设电子质量为 m，电子绕核圆周运动速度为 v，r 为电子可能的轨道半径。根据经典力学，电子受到的向心力等于质子与电子的库仑引力，即

$$\frac{e^2}{4\pi\varepsilon_0 r^2} = \frac{mv^2}{r} \tag{1-16}$$

式中，ε_0为真空介电常数；e为电子电量。将玻尔角动量量子化条件式（1-15）代入式（1-16），消去两式中的v，并以r_n代替r，得

$$r_n = \frac{4\pi\varepsilon_0 n^2\hbar^2}{me^2} \quad (n=1, 2, 3, \cdots) \tag{1-17}$$

这就是氢原子中第n个稳定轨道半径。当n=1时，是氢原子电子的最小轨道半径，称为玻尔半径a_0。由式（1-17）求得$a_0 = 5.29\times10^{-11}$m，这个值的数量级与实验数据一致。

当电子在量子数为n的轨道上运动时，原子系统的总能量E_n等于电子的动能$\frac{1}{2}mv_n{}^2$和电子与原子核系统的势能$-\frac{e^2}{4\pi\varepsilon_0 r_n}$的代数和，即

$$E_n = \frac{1}{2}mv_n{}^2 - \frac{e^2}{4\pi\varepsilon_0 r_n} \tag{1-18}$$

将式（1-16）和式（1-17）代入式（1-18）最后得

$$E_n = -\frac{me^4}{8\varepsilon_0{}^2 n^2 h^2} \quad (n=1, 2, 3, \cdots) \tag{1-19}$$

式（1-19）表示电子在第n个稳定状态中原子系统的能量。这种量子化的能量值称为原子能级（简称能级）。

按照玻尔第二个假设，原子系统中电子从较高能级E_n跃迁到较低能级E_k时，所发射的单色光的波数$\tilde{v}_{kn}$为

$$\tilde{v}_{kn} = \frac{v_{kn}}{c} = \frac{me^4}{8\varepsilon_0{}^2 h^3 c}\left(\frac{1}{k^2} - \frac{1}{n^2}\right) \tag{1-20}$$

式中，c为光速。式（1-20）与氢原子光谱的经验公式（1-14）是一致的。由式（1-20）计算的里德伯常数R_H与实验值十分符合。图1-7和图1-8中均示出了氢原子能态跃迁所产生的各种谱线系。

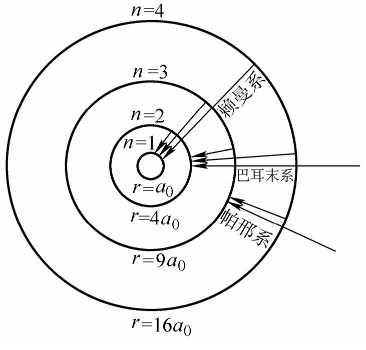


图1-7　氢原子的电子轨道及光谱线

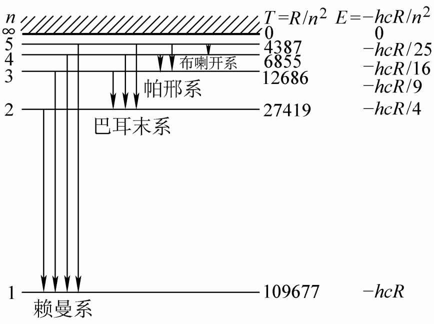


图1-8　氢原子能级

（5）玻尔理论成功和缺陷

玻尔理论不仅能成功地说明氢原子的光谱，对类氢离子（只有一个电子绕核转动的离子）的光谱也能很好地说明。由此可见，玻尔理论在一定程度上能反映单电子原子系统的客观实际。鉴于玻尔对研究原子结构和原子辐射的贡献，玻尔荣获 1922 年诺贝尔物理学奖。同时玻尔关于定态的概念和光谱线频率的假设，在原子结构和分子结构的现代理论中，仍然是有用的概念。玻尔的创造性工作对现代量子力学的建立有着深远的影响。

玻尔理论虽然取得一些成就，但是也存在着严重不足之处。首先，这个理论本身仍是以经典理论为基础的，而所引进的电子处于定态时不发出辐射的假设却又是和经典理论相抵触的。其次，量子化条件的引进也没有适当的理论解释。此外，由玻尔理论只能求出谱线的频率，对谱线的强度、宽度、偏振等一系列问题都无法处理。

1.1.2 德布罗意物质波

以玻尔理论为基础的量子理论（称之为旧量子论），尽管取得了不少令人惊奇的成果，但也存在着严重的缺陷与不足，遇到了越来越多的困难。1923 年，德布罗意提出“物质具有波粒二象性”的假设，创造性地做出“电子具有波动性”的预言，为量子论的发展开辟了一条崭新的途径，打通了薛定谔通往波动力学的道路。1926 年，玻恩提出波函数的统计解释，找到了波与粒子统一的线索，把波与粒子统一到概率波概念的基础上，为量子力学提供了全新的概念基础。1927 年，海森伯提出了不确定性关系（或原理），加深了人们对量子本质的认识，极大地推进了人类认识微观客体本质属性的历史进程。

1. 德布罗意物质波假设

在汤姆孙通过研究阴极射线确定射线的荷质比并确定电子的存在时，人们认识到电子的微粒特性。至于电子的波动性则是在发现电子 27 年后由年仅 25 岁的德布罗意大胆假设实物微粒也应像光子一样具有波粒二象性之后提出的。德布罗意后来谈到他如何把光的波粒二象性推广到实物粒子上来的。他在 1964 年致库伯利（Kubli）的信中写道：在 1922～1923 年期间，当我开始获得波动力学的基本想法时，我的意图是把爱因斯坦发现的光的波粒共存现象推广到所有粒子，因此，我开始把爱因斯坦建立的光量子公式 $\varepsilon = h\nu$ 和 $p = h\nu/c = h/\lambda$，应用到光子以外的其他粒子上。

德布罗意认为，质量为 m 的粒子、以速度 v 匀速运动时，具有能量 E 和动量 p；从波动性方面来看，它具有波长 λ 和频率 ν，而这些量之间的关系也和光波的波长、频率与光子的能量、动量之间的关系一样，应遵从下述公式

$$E = mc^2 = h\nu \tag{1-21}$$

$$p = mv = \frac{h}{\lambda} \tag{1-22}$$

所以对具有静止质量 m_0 的实物粒子来说，若粒子以速度 v 运动，则和该粒子相联系的平面单色波的波长 λ 是

$$\lambda = \frac{h}{p} = \frac{h}{mv} = \frac{h}{m_0 v}\sqrt{1-\frac{v^2}{c^2}} \tag{1-23a}$$

式（1-23）称为德布罗意公式。人们常把这种和物质相联系的波通常称为德布罗意波。薛定谔在诠释波函数的物理意义时，把这种波称为物质波。如果 $v \ll c$，那么

$$\lambda = \frac{h}{m_0 v} \tag{1-23b}$$

德布罗意把他的相波概念应用到以闭合轨道绕核运动的电子，从而推出了玻尔的量子化条件，这是德布罗意波能够成立的一个有力证据。德布罗意提出“只有满足相波谐振的那条轨道才是稳定的轨道。”电子的轨道是由 $\frac{\nu}{v}$ 对闭合路径的积分取整数值来确定，即

$$\oint \frac{\nu}{v}\mathrm{d}l = n \quad (n=1, 2, 3, \cdots) \tag{1-24}$$

谐振的条件是 $l=n\lambda$，即电子轨道的周长是相波波长的整数倍，亦即

$$2\pi r = n\lambda = n\frac{h}{mv} \tag{1-25}$$

这正是玻尔假设中有关电子轨道角动量量子化的条件。这就是说，氢原子中作稳定的圆周运动的电子相应的波必须是一个驻波，其形状如图 1-9 所示。即电子绕核一周后，驻波应光滑地衔接起来。

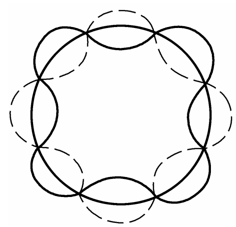


图 1-9　电子的德布罗意波

2. 物质波的实验验证

最早从实验上证实电子衍射现象的是美国的戴维孙（C. J. Davisson）和他的合作者革末（L. H. Germer）。1926 年他们将电子枪射出的电子束投射到镍单晶体表面上，研究了散射电流与轰击电压和散射角的关系，肯定了这是电子衍射的结果。经过定量计算，证明了德布罗意波长公式的正确性。

图 1-10 是戴维孙和革末的电子在晶体上衍射的实验装置示意图。从加热灯丝 K 出来的电子，经电势差 U 加速后，通过一组阑缝 D，成为很细的平行电子射线，以一定的角度投射到镍单晶体 M 上，经晶面反射后用集电器 B 收集。进入集电器的电子流强度 I 可用与 B 相连的电流计 G 来量度。实验时，使图中所示的两个 φ 角相等，并保持不变；改变电势差 U，量度相应的电子流强度 I。实验结果如图

1-11 所示，在电势差 U 单调地增大时，电子流强度 I 不是单调地增大，只有当电势差为某些特定值时，电子流才有极大值。

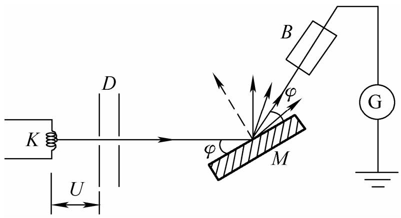


图 1-10　电子在晶体上衍射实验装置示意图

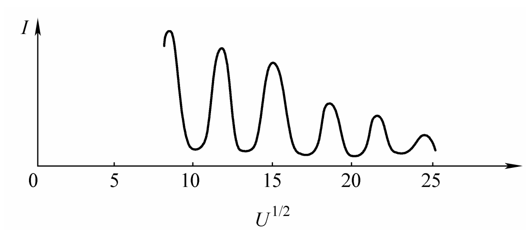


图 1-11　电子流强度与电势差的关系

为了解释这一实验结果，用 X 射线衍射的结果来分析。X 射线在晶体上反射时，只有入射线的波长 λ 符合布拉格公式 $2d\sin\varphi = k\lambda$）$k = 0,1,2,\cdots$；d 为晶面间距）的那些射线才能在一定的角度 φ 观察到反射线。如果承认电子也有波的性质，并把上述实验结果看做是电子波的衍射。那么，在 d 和 φ 给定时，仅当电子波波长 λ 满足布拉格公式时，才能按反射定律自晶面反射。由德布罗意波公式和布拉格公式得

$$2d\sin\varphi = kh\sqrt{\frac{1}{2em_0}}\frac{1}{\sqrt{U}} \tag{1-26}$$

即电势差 U 满足式（1-26）时，电子流强度 I 为最大值。由此计算所得的电势差 U 的各个量值和实验结果相符合，从而证明了德布罗意物质波假设是正确的。

1927 年，汤姆孙（G. P. Thomson）和里德做了高能电子束透射金属薄层的实验，观察到了衍射环。根据这些圆环的半径计算出电子波的波长，证明了德布罗意公式，证实了电子的波粒二象性。1960 年约恩孙（C. Jonson）直接做了电子双缝干涉实验，从屏上得到了类似光的杨氏双缝干涉图样。电子的波动性获得了实验证实以后，在其他的一些实验中也观察到中性粒子，如原子、分子和中子等微观粒子，也具有波动性，同样证实德布罗意公式的正确性。由此可见，一切微观粒子都具有波动性，德布罗意波的存在已是确实无疑的了。德布罗意公式已成为

揭示微观粒子的波—粒二象性的统一性的基本公式。

3. 物质波是概率波

（1）玻恩的统计解释

1926 年，波恩提出了德布罗意波的统计意义，认为波函数体现了发现粒子的概率，这是每个粒子在它所处环境中所具有的性质。如果有大量的粒子，那么在某处粒子的密度就与此处发现一个粒子的概率成正比。把这情况同光来对比，光的强弱同光子的数目成正比，而在某处的光子数同该处发现一个光子的概率成正比，但我们知道光的强弱是同光波的电场或磁场强度的平方成正比的，即与光波的波函数的平方成正比，足见在某处发现一个光子的概率同光波的电场或磁场强度的平方（即波函数的平方）成正比。这样来类比，可见在某处发现一个实物粒子的概率同德布罗意波的波函数平方 ψ^2 成正比。可以把体积 $\mathrm{d}\tau$ 中发现一个粒子的概率表达为 $\mathrm{d}\omega=\psi^2\mathrm{d}\tau$。由此，$\psi^2$ 代表在单位体积内发现一个粒子的概率，因而称概率密度。这就是德布罗意波函数的物理意义。

玻恩提出的波函数 ψ 与经典的波函数有类似之处，但是它们的意义是完全不同的。经典的波函数意味着有某种实在的物理量的空间分布作周期性的变化，是可以测量的，而概率波 ψ 在一般情况下是不可测量的。一般可以测量的只是 ψ^2，它的含义是概率。对于概率分布来说，重要的是相对概率分布。显而易见，$\psi(r)$ 与 $C\psi(r)$（C 为常数）所描述的相对概率分布是完全相同的；而经典波的波幅若增加一倍，则相应的波动能量将为原来的 4 倍，代表完全不同的波动状态。

（2）概率波对电子双缝干涉实验的解释

怎样解释电子双缝干涉实验？是否可以用波的干涉理论来说明？电子运动的波列到达双缝（双缝假设为 A 和 B）后，将在双缝 A 和 B 处形成两个子波，两个波叠加后在接收屏上得到双缝干涉图样。但是这样的描述与电子粒子性相矛盾。到目前为止，我们知道电子是不可分割的，一个电子通过 A 和 B 时被分成两部分，彼此进行干涉是不可能的。反之，如果认为电子到达双缝时只有一条狭缝起作用，即每个电子仅仅从其中一条通过，则这种观点虽然维护了电子的粒子性，却又与实验事实明显不符。应该说，电子通过缝 A 时，缝 B 是否打开对它不应有任何影响。同样，电子通过缝 B 时，缝 A 的存在也应与它毫无关系。按照这种观点必然得到如下结论：先打开缝 A 关上缝 B，然后再打开缝 B 关上缝 A，应得到与同时打开双缝相同的衍射图样。但事实却不是如此，先后依次打开一条缝得到的必然是两个单缝衍射图样的叠加，屏上电子的强度应是分别打开时强度之和，不可能得到双缝干涉图样。这就是说，尽管粒子性要求每个电子只可能从一条缝通过，但双缝仍然同时在起作用，与只打开单缝电子的“感受”是不同的。

概率波的概念既能维护电子的粒子性，又能体现电子的波动性。由于概率波给出的是电子出现的概率，所以不要求分割电子。而电子的概率分布又是由波函

数决定的，所以能给出干涉、衍射图样，又体现了波动性。电子通过双缝时，同时打开两个缝与分别先后各打开一个缝对于电子来说是有区别的，在前一情形每个电子都有两种机会，即可以从 A 或 B 通过，而在后一情形就只存在一种机会。因此，应当分别用两种不同的概率波来表示这种情况，即波函数 ψ 应是两种不同的数学表达式，分别给出不同的干涉图形，即分别是双缝或单缝衍射图样。一个表示有机会通过 A 和 B 两个缝的概率波并不破坏电子的粒子性，它是说明每个电子有机会通过缝 A 也有机会通过缝 B，而没有要求一个电子分成两半，分别通过 A 和 B。

电子到达接收屏后也不会发生矛盾，这时经过衍射的概率波将给出电子在屏上分布的概率，例如亮条纹处概率大，暗条纹处概率小等。电子到达屏时实际上只落在一个点上，不过落在哪个点上是按概率波给出的概率分布的。在大量电子参加双缝实验时，各个电子都以同样的概率波表示，大量电子到达接收屏后必然在概率大的地方电子多（出现亮条纹），在概率小的地方少（出现暗条纹）。因此，用概率波描述电子能很好地将电子的波动性和粒子性结合起来。

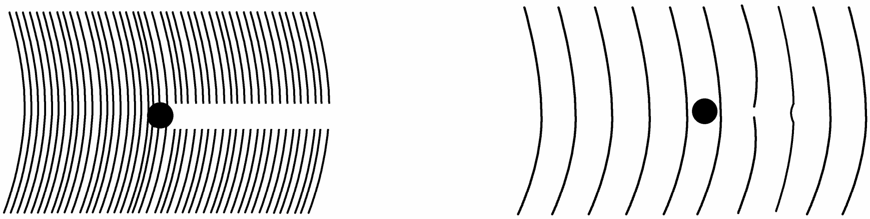
1.1.3 测量和测不准原理

1. 宏观物体的测量和微观物体的测量

当研究物体的运动时，最初必须建立观测运动的手段。也就是说，应该跟踪它的轨道。使用可见光或电磁波跟踪宏观物体运动的理由，除了用这种方法传递信息非常快的优点之外，还由于光照射到物体上而物体的运动几乎没有改变的缘故。可见光的光子动量为 10^{-27}kg·m·s^{-1} 数量级，这样大小的动量与宏观物体运动时的动量（如 1g 的物体具有 10m·s^{-1} 速度时的动量为 10^{-2}kg·m·s^{-1}）相比要小的多，所以光触及在物体上而产生的影响是完全可以忽视的。然而，当物体的质量非常小而其动量和光的动量具有相同大小时，物体的运动状态则由于光的触及发生变化。但是如果光没有触及到物体，就不可能知道它的位置，当然，也得不到关于运动状态的数据。

为了使物体受到的影响小，是否可用动量小而波长长的电磁波呢？可是在这种情况下，虽然动量变小，但由于波长长而给确定物体的位置带来不准确性。图 1-12a、b 分别表示用短波和长波测量物体时的情况。但是在后者的情况下，由于波的衍射效应，物体不能阻止波的传播，因此，不能起到测量位置的作用。

实际上，要这样研究物体的运动时，必须要考虑观测问题。可是在牛顿力学中，认为这个问题是与运动定律无关系的次要问题。确实在研究一般物体运动时，观测只是手段。当进行观测的时候，对运动状态本身不会有任何影响。将这种观点无限制地推广，为了观测，所进行的操作对物体运动状态产生很大影响时，就会认为牛顿力学仍旧适用的看法是有问题的。



a)　　　　b)

图 1-12　物体对波的影响

a）短波情况　b）长波情况

运动状态不受影响而能跟踪电子的位置是不可能的。像阴极射线管（示波器），电子束碰撞在荧光膜上而发出荧光，由于能发光的荧光屏可以表示自身的位置，所以通过管内的电子在这一瞬间，电子的动量与碰撞前是完全不同的。再有关于荧光屏上的电子位置，严密地来说，小于荧光光点大小的是不能正确指示的。又因为电子束是由许多电子组成的。究竟是哪个电子在什么时候碰撞在荧光屏上，要正确地获得这些数据是格外困难的。

2. 测不准原理

测不准关系是海森伯在1927年根据对一些理想实验的分析和德布罗意关系而得出的。即微观粒子在位置与动量两者不确定量之间的关系满足

$$\Delta x\Delta p_x \geqslant h，\quad \Delta y\Delta p_y \geqslant h，\quad \Delta z\Delta p_z \geqslant h \tag{1-27}$$

最初，海森伯用波包概念第一次定量地导出了粒子坐标和动量之间的不确定关系。当玻尔看到论文的初稿时，认为波粒二象性应当成为量子理论解释的出发点，并且果然通过波粒二象性的论证，重新导出了不确定关系。玻尔说，一个在空间和时间上都是有限的实际波场，只能由原则上与无限多的频率 v 值和波矢 k 值相对应的基本波叠加而成。他从经典波动概念出发，把物质粒子的波动性看成是三维空间中连续分布的波包（图 1-13），对于一个位置确定的波包，必须满足 $\Delta x\Delta\lambda\geqslant\lambda^2$。即要无限精确地测准波长（$\Delta\lambda\to0$），就必须在无限扩展的空间中进行观测（$\Delta x\to\infty$）。再运用德布罗意关系 $\lambda = h/p$ 得到 $\Delta\lambda = (h/p_x^2)\ \Delta p_x$ 。代入 $\Delta x\Delta\lambda\geqslant\lambda^2$ 式，便有

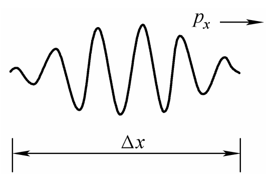


图 1-13　波包

$$\Delta x\Delta p_x \geqslant h \tag{1-28a}$$

这就是粒子位置和动量的不确定关系式。

海森伯提出不确定关系时，量子力学的基本数学结构和物理解释都已建立起来。1927 年海森伯的文章刚发表出来，Kennard 就运用量子力学的数学公式推导

出不确定关系

$$\Delta x \Delta p_x \geqslant \frac{\hbar}{2}, \quad \Delta E \Delta t \geqslant \frac{\hbar}{2} \tag{1-28b}$$

式中，ΔE 表示能量的不确定量；Δt 表示时间的不确定量。

下面利用波长很短的 γ 射线显微镜的设想实验对这个问题作详细的讨论。图 1-14 所示的是用 γ 射线确定电子位置的装置。入射的 γ 射线的光子和电子碰撞，并通过透镜应该到达检测器。因为 γ 射线能通过 OAP，也应该通过 OBP，所以，∠AOB 为 2α 时，光子动量的 x 分量具有 $\frac{h\nu}{c}\sin\alpha$ 和 $-\frac{h\nu}{c}\sin\alpha$ 的宽度。设它为 Δp_x 就成为

$$\Delta p_x = 2\frac{h\nu}{c}\sin\alpha \tag{1-29}$$

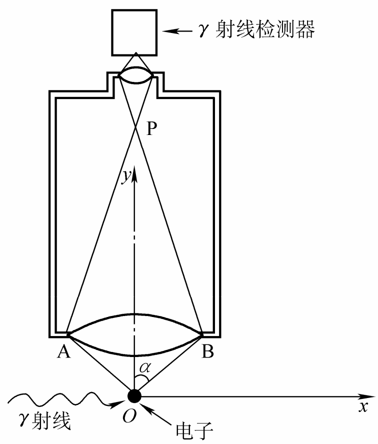


图 1-14 γ 射线显微镜（设想实验）

另一方面，根据物理光学，显微镜的分辨率或者 x 位置的不确定性，由下式给出

$$\Delta x \approx \frac{c}{\nu \sin\alpha} \tag{1-30}$$

由式（1-29）和式（1-30）得如下关系

$$\Delta x \Delta p_x \approx h \tag{1-31}$$

在实际上，因为不确定性更大，所以可写为

$$\Delta x \Delta p_x \geq h \tag{1-32}$$

在图 1-14 中，用 γ 射线捕获电子时，假设只有 Δx 的不确定性，这意味着在 γ 射线捕获的时间 t 内，只有如下的不确定性

$$\Delta t = \frac{\Delta x}{v_x} \tag{1-33}$$

式中，v_x 是 γ 射线碰撞前电子速度的 x 分量。另一方面，因为电子的能量为 $E = \frac{1}{2m}({p_x}^2 + {p_y}^2 + {p_z}^2)$，所以当 p_x 只有 Δp_x 的变化时，则 E 的变化为

$$\Delta E = \left(\frac{\partial E}{\partial p_x}\right)\Delta p_x = \frac{p_x}{m}\Delta p_x = v_x \Delta p_x \tag{1-34}$$

由式（1-32）、式（1-33）和式（1-34）得下列关系

$$\Delta E \Delta t = \Delta p_x \Delta t \geqslant h \tag{1-35}$$

也就是说，电子的能量和在它观测所需要的时间 t 内，可以有不确定性或是互补的关系。

3. 测不准关系的物理解释

在不确定关系提出之后，物理学家对它的物理本质具有不同理解，对它是由什么因素引起以及它的意义是如何具有不同评述。

海森伯和波尔等人提出了一种“扰动说”。认为不确定关系来源于在单次测量中宏观仪器对被测量的微观系统的不可控制和不可预计的扰动。在海森伯最早关于 γ 射线显微镜的论证里，他认为由于“光子被电子散射的瞬间，电子的动量产生一个不连续的改变”，“因此，电子的位置测定得愈精确，动量就知道得愈不精确，反之亦然”。玻尔也讲过：“按照量子理论，正是由于不可能忽略系统和测量器械的相互作用，就意味着每一次观察都将引入一个新的不可控制的要素。”

另一种看法认为不确定关系是波粒二象性的反映。玻尔就是根据粒子的波粒二象性来说明不确定关系的。其方法是假定电子完全可用波概念来表示，那么粒子概念可应用性就受到限制，这个限制就由海森伯关系表达出来。换句话说，假定波理论的最简单的概念是对的，这些波概念就可取来作为评判粒子图景的基础，从而显示出粒子图景仅在一定限度上可用。例如，如果一个粒子出现的地点已被确切地确定了，则用波图景就必须说：在这里存在着一个缩拢到一个很小空间上的“波包”。可是这样一个波包是不稳定的，它立刻分裂散开。但是用粒子图景，这就意味着对这粒子此后运动的速度不能赋予任何确定的数值和方向。在这里应用粒子速度概念在一定程度上是不确定的，事实上粒子位置知道得越准，粒子速度就越不确定。

在玻恩提出波函数的统计解释以及用量子力学的基本原理导出不确定关系后，物理学家提出对不确定关系的统计解释。即不确定关系的不确定量都应当指的是多次测量里观测值偏离其平均值的统计散布，不确定关系就是两种统计散布（即涨落）的乘积所受到的限制。

现在的物理教科书一般都主张不确定关系是微观粒子波粒二象性的宏观表现，认为粒子的波动性使得粒子在任何一时刻不具有确定的位置，也不具有确定的动量。

不确定关系说明了用经典力学描述微观粒子运动时所存在的局限性，给人们指出了使用经典粒子运动概念的一个限度。不确定关系划分了经典力学和量子力学的界限。

1.1.4　薛定谔方程

薛定谔把玻尔原子系统的能级理论、德布罗意物质波思想以及哈密顿的经典

力学与几何光学的数学相似性思想结合起来，形成了波动力学概念，建立了薛定谔方程。方程建立过程中一个突出特点就是善于运用物理类比。薛定谔通过力学与光学类比形成波动力学概念，通过几何光学与经典力学的相似性得出波动光学与波动力学的相似性，从而根据普遍的波动方程建立起了薛定谔方程。

1. 波函数

具有能量 E 和动量 p 的自由运动的一个微观粒子必然表现出波动性，我们用波函数来描述它的运动状态。自由粒子不受力，动量 p 不变，能量 E 不变，所以同它联系的波长 λ 不变，频率 v 不变，波的传播方向也不变，是平面单色波。我们知道，频率为 v、波长为 λ 沿 x 轴方向传播的平面单色波的可以用下式表示

$$\psi(x,t)=\psi_0\cos 2\pi\left(vt-\frac{x}{\lambda}\right) \tag{1-36a}$$

将这个公式写成复数形式

$$\psi(x,t)=\psi_0 \mathrm{e}^{-\mathrm{i}2\pi\left(vt-\frac{x}{\lambda}\right)} \tag{1-36b}$$

而只取其实部。将式（1-21）和式（1-22）代入式（1-36b）便得到自由粒子的平面波

$$\psi(x,t)=\psi_0 \mathrm{e}^{-\mathrm{i}\frac{2\pi}{h}(Et-px)}=\psi_0 \mathrm{e}^{\frac{\mathrm{i}}{\hbar}(px-Et)} \tag{1-37}$$

如果粒子在三维空间运动，则上式可推广为

$$\psi(\boldsymbol{r},t)=\psi_0 \mathrm{e}^{\frac{\mathrm{i}}{\hbar}(\boldsymbol{pr}-Et)} \tag{1-38}$$

$\boldsymbol{r}$ 是原点到波面任何一点的径矢。

玻恩对波函数的统计解释：波函数（或波函数振幅）的平方表示在单位体积中发现一个粒子的概率，称为概率密度。波函数是复数，而概率却必须是实正数，所以波函数的平方，应等于波函数 ψ 与其共轭复数 ψ^*的乘积。波函数既然具有这样的物理意义，必须满足一些条件。它必须是连续的、单值的、有限的，概率不会在某处发生突变，所以函数必须随处连续。在任何处，只能有一个概率，所以函数在任何处是单值的。概率不能无限大，所以函数值必须有限。不符合这三个条件的波函数是没有物理意义的，它不代表物理实在。另外，由于粒子必定要在空间中的某一点出现，所以粒子在空间各点出现的概率总和等于 1，也就是

$$\iiint|\psi|^2\mathrm{d}V=1 \tag{1-39}$$

式中，V 为空间体积。式（1-39）称为归一化条件，其积分应遍及整个空间。

2. 定态薛定谔方程

一个微观粒子的状态可用波函数 $\psi(r, t)$来描述，当 $\psi(r, t)$确定以后，在空间某

处发现一个粒子的概率以及任何力学量的测量值概率都完全确定了。因此，量子力学中最核心的问题就是要找出各种具体情况下描述系统状态的各种可能的波函数。这个问题由薛定谔提出的波动方程得以圆满解决。薛定谔方程是量子力学的基本方程，它不能由其他原理推导出来，它的正确性只能靠实验来检验。

下面是建立定态薛定谔方程的主要思路（并不是方程的理论推导）。

一个具有能量 E 和动量 p 的自由粒子波函数由式（1-38）给出，它可以写为

$$\psi(\boldsymbol{r},t)=\psi(\boldsymbol{r})\mathrm{e}^{-\frac{\mathrm{i}}{\hbar}Et} \tag{1-40}$$

式中，

$$\psi(\boldsymbol{r})=\psi_0\mathrm{e}^{\frac{\mathrm{i}}{\hbar}\boldsymbol{pr}} \tag{1-41}$$

$\psi(\boldsymbol{r})$ 称为振幅函数，它是波函数中只与坐标有关，而与时间无关的部分，也称为波函数。振幅函数分别对 x、y、z 求二阶导数，将求得导数的三式相加，即有

$$\nabla^2\psi=-\frac{p^2}{\hbar^2}\psi \tag{1-42}$$

其中 ∇^2 为拉普拉斯算符，定义为

$$\nabla^2=\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial x^2} \tag{1-43}$$

如果粒子不是自由粒子而是在势场中运动，上述方程仍然成立。粒子的总能量 E 应是势能 U 和动能 $\frac{1}{2}m\upsilon^2$ 之和，那么有 $p^2=2m(E-U)$，将其代入（1-42）得

$$-\frac{\hbar^2}{2m}\nabla^2\psi+U\psi=E\psi \tag{1-44}$$

ψ 只是坐标的函数，式（1-44）不含时间，这就是定态薛定谔方程。

3. 薛定谔方程的意义

薛定谔方程是量子力学的基本方程，它揭示了微观物理世界物质运动的基本规律。它是原子物理学中处理一切问题的有力工具，在原子、分子、固体物理、核物理、化学等领域被广泛应用。

薛定谔认为，玻尔等人的量子论都是假设的，认为玻尔的这些“量子条件具有完全奇怪的和不可理解的性质”。应该取代为在满足波函数单值、有限、连续的条件下以及一定的边界条件下求解波动方程的本征值，从而求得一组分立的本征率。“整数性质是这一方法的自然结果”，而不是人为规定。薛定谔方程的解确实具有整数性，它真正地将量子论带到科学的领域。

狄拉克把狭义相对论引入到量子力学中，这就成为相对论的量子力学的基础。他还运用众所周知的狄拉克符号法这种独特的表述方法，发展了称为变换理论的一般的量子理论，由此，量子论的体系可以说基本完成。

1.2 波动现象

1.2.1 波的基本方程式

在自然界里有声波、弹性波、电磁波等各种波动现象。所谓波动，就是以介质中的各点作为中心，振动在它的平衡位置上而在空间传播的现象。在广义的波动现象中，变化的量没有必要由平衡位置的变位来限定，一个物理量如果能随时间、空间变化，就是一种信息在空间的传播现象。因此，热传导现象也可以认为是波动现象。

假如 ψ 看作物理量时，给出 ψ 的空间的、时间的变化的微分方程式，叫作波动方程式。研究一下细弦的横向振动，就是一个简单的例子。如图 1-15 所示，试分析一下这个弦的很小一部分 AB。设弦的张力为 T，张力 T 在 A 点中的 y 分量为$-T\sin\theta$，而 B 点是 $T\sin\theta'$。如果认为 θ、θ'是很小的话，则用于 AB 部分的 y 分量合力 dF 为

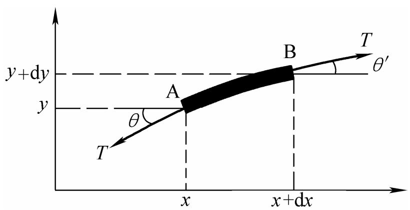


图 1-15 弦的振动

$$\mathrm{d}F = T(\sin\theta' - \sin\theta) \approx T(\tan\theta' - \tan\theta) \tag{1-45}$$

因为

$$\tan\theta = \left(\frac{\partial y}{\partial x}\right)_x ;\quad \tan\theta' = \left(\frac{\partial y}{\partial x}\right)_{x+\mathrm{d}x} = \left(\frac{\partial y}{\partial x}\right)_x + \left(\frac{\partial^2 y}{\partial x^2}\right)\mathrm{d}x \tag{1-46}$$

所以，由式（1-45）和式（1-46）可得

$$\mathrm{d}F = T\left(\frac{\partial^2 y}{\partial x^2}\right)_x \mathrm{d}x \tag{1-47}$$

另一方面，AB 部分的 y 方向加速度，如果用 t 表示时间，则以 $\dfrac{\partial^2 y}{\partial t^2}$ 表示。设 σ 为弦的线密度时，质量为 $\sigma\mathrm{d}x$，去掉 $\left(\dfrac{\partial^2 y}{\partial x^2}\right)_x$ 的角标 x，根据牛顿第二定律，则有

$$\sigma\mathrm{d}x\frac{\partial^2 y}{\partial t^2} = \mathrm{d}F = T\frac{\partial^2 y}{\partial x^2}\mathrm{d}x$$

即

$$\frac{\partial^2 y}{\partial x^2} = \frac{\sigma}{T}\frac{\partial^2 y}{\partial t^2} \tag{1-48}$$

这是关于弦横向变位 y 的时间 t 和 x 坐标的二阶偏微分方程式，这就是波动方程式。

以上是弦的横向振动，在一般情况下，波动方程式稍微复杂一点，如下面的二阶偏微分方程式

$$\frac{\partial^2\Psi}{\partial x^2}+g(x)\Psi=a\frac{\partial\Psi}{\partial t}+b\frac{\partial^2\Psi}{\partial x^2} \tag{1-49}$$

式中 Ψ 表示变化的物理量，是波函数的总称。如果 Ψ 是 $\Psi(x, y, z, t)$的三维变化量时，波动方程式为

$$\frac{\partial^2\Psi}{\partial x^2}+\frac{\partial^2\Psi}{\partial y^2}+\frac{\partial^2\Psi}{\partial z^2}+g(x,y,z)\Psi=a\frac{\partial\Psi}{\partial t}+b\frac{\partial^2\Psi}{\partial t^2} \tag{1-50}$$

或者用下式表示

$$\nabla^2\Psi+g(x,y,z)=a\frac{\partial\Psi}{\partial t}+b\frac{\partial^2\Psi}{\partial t^2} \tag{1-51}$$

∇^2 是拉普拉斯算符。$g(x, y, z)$是在空间能够变化的系数；a、b 是常数。例如，在麦克斯韦的电磁方程式中，设 ε、μ、κ 分别为介电常数、磁导率、电导率，则电场矢量 $\boldsymbol{E}$、磁场矢量 $\boldsymbol{H}$ 分别要满足波动方程式

$$\nabla^2E=\kappa\mu\frac{\partial E}{\partial t}+\varepsilon\mu\frac{\partial^2E}{\partial t^2}；\quad \nabla^2H=\kappa\mu\frac{\partial H}{\partial t}+\varepsilon\mu\frac{\partial^2H}{\partial t^2} \tag{1-52}$$

在这种情况下，Ψ 为 E 或 H；$g=0$；$a=\kappa\mu$；$b=\varepsilon\mu$。

根据式（1-50）或者式（1-51）完全可以求出波动现象，但得到什么样的解？这取决于 Ψ 的边界条件和起始条件。

1.2.2　谐波

在波当中最简单而且基本的波，可以说是谐波。它是在均匀介质中向一定方向进行的正弦波。在空间变化的物理量是实数，但作为数学上的方法则利用复函数表示更方便。谐波形式为

$$\Psi=Ae^{\mathrm{i}(kx-\omega t)}=Ae^{2\pi\mathrm{i}(x/\lambda-\nu t)} \tag{1-53}$$

式中 A 是常数，一般为复数，如 $A=|A|\mathrm{e}^{-\mathrm{i}\delta}$ 时，式（1-53）就成为

$$\Psi=|A|\mathrm{e}^{\mathrm{i}(kx-\omega t-\delta)} \tag{1-54}$$

实际的波是用 Ψ 的实数部分 Ψ_r 或者虚数部分的系数 Ψ_i 来表示。即

$$\left.\begin{aligned}\Psi_r&=|A|\cos(kx-\omega t-\delta)\\ \Psi_i&=|A|\sin(kx-\omega t-\delta)\end{aligned}\right\} \tag{1-55}$$

在上式中的常数分别为：$|A|$振幅；k 波数（$=2\pi/\lambda$）；λ 波长；ν 频率；ω 角频率；δ

位相因子。

由式（1-55）表示的波，显而易见是在 x 方向传播的平面波。Ψ_r 和 Ψ_i 如图 1-16 所示，波在空间具有 λ 的周期，在 x 方向上，以 u_p 的速度前进，u_p 为

$$u_p = \frac{\lambda}{T} = \lambda\nu = \frac{\omega}{k} \tag{1-56}$$

式中 T 为周期。u_p 叫做波的传播速度或者叫相速。

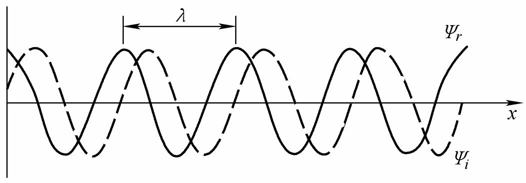


图 1-16 谐波

谐波在任意方向传播时，如果把波数用波数矢量（简称波失）$\boldsymbol{k}(k_x, k_y, k_z)$替换得

$$\Psi = A\mathrm{e}^{\mathrm{i}(\boldsymbol{kr}-\omega t)} = A\mathrm{e}^{\mathrm{i}(k_x x+k_y y+k_z z-\omega t)} \tag{1-57}$$

式中 $\boldsymbol{r}$ 是 x、y、z 分量的位矢。在对某时间取 $\boldsymbol{kr}$ 一定值的面中，因为 Ψ 成为同位相，所以这个面应该成为波面，可是因为 $\boldsymbol{k}$ 矢量与该面正交，所以波可以说是在 $\boldsymbol{k}$ 矢量的方向上传播。

在波动现象中，振动发生在波的传播方向的波叫纵波，振动发生在波传播方向的垂直方向的波叫做横波，如图 1-17 所示。在这样的情况下，式（1-57）中的 A 被 A 的矢量所代替。

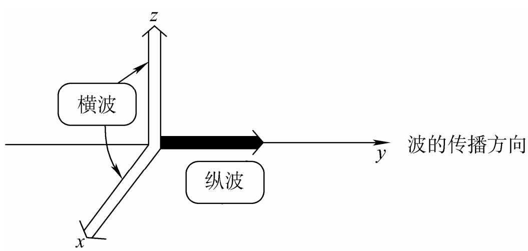


图 1-17 波的传播方向与振动方向

1.2.3 变数分离

讨论波函数 Ψ 为如下形式的情况

$$\Psi = \psi(x)\phi(t) \tag{1-58}$$

也就是看做是空间坐标函数$\psi(x)$和时间函数$\phi(t)$的积。如果将式（1-58）代入式（1-49）则得

$$\frac{1}{\psi}\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2}+g(x)=\frac{a}{\phi}\frac{\mathrm{d}\phi}{\mathrm{d}t}+\frac{b}{\phi}\frac{\mathrm{d}^2\phi}{\mathrm{d}t^2} \tag{1-59}$$

因为式（1-59）的左边只是 x 的函数，右边只是 t 的函数，因此，它们必须等于常数。如果把这个常数称为 K 时，式（1-59）就成为如下形式

$$\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2}+[g(x)-K]\psi=0 \tag{1-60}$$

$$b\frac{\mathrm{d}^2\phi}{\mathrm{d}t^2}+a\frac{\mathrm{d}\phi}{\mathrm{d}t}-K\phi=0 \tag{1-61}$$

这样就分离成 x 和 t 的二阶全微分方程式。这种操作叫做变数分离。

假如，波函数由式（1-53）给出时，可写成如下的关系

$$\psi=A\mathrm{e}^{\mathrm{i}kx},\ \ \phi=\mathrm{e}^{-\mathrm{i}\omega t} \tag{1-62}$$

将式（1-62）代入式（1-60）和式（1-61）就得到

$$-k^2+g(x)=K \tag{1-63}$$

$$-\omega^2 b-ia\omega=K \tag{1-64}$$

因为 K 是常数，所以在式（1-63）中的 $g(x)$也必须是常数。并且式（1-63）和式（1-64）给出了波数 k 和角频率 ω 的关系。

要解 x 的微分方程式（1-60），当利用式（1-63）时，式（1-60）可写成

$$\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2}+k^2\psi=0 \tag{1-65}$$

k 为常数，它的一般解为

$$\psi=A_+\mathrm{e}^{\mathrm{i}kx}+A_-\mathrm{e}^{-\mathrm{i}kx} \tag{1-66}$$

式中 A_+、A_-是边界条件决定的常数。

现在来看一个边界条件为$\psi(x_0/2)=\psi(-x_0/2)=0$ 求解$\psi(x)$的特例，这里 x_0 为某一具体值。由边界条件可得

$$\left.\begin{aligned}A_+\mathrm{e}^{\frac{1}{2}\mathrm{i}kx_0}+A_-\mathrm{e}^{-\frac{1}{2}\mathrm{i}kx_0}=0\\A_+\mathrm{e}^{-\frac{1}{2}\mathrm{i}kx_0}+A_-\mathrm{e}^{\frac{1}{2}\mathrm{i}kx_0}=0\end{aligned}\right\} \tag{1-67}$$

这是关于 A_+、A_-的联立齐次方程组。若得非零解，系数的行列式必须为零。即

$$\begin{vmatrix}\mathrm{e}^{\frac{1}{2}\mathrm{i}kx_0} & \mathrm{e}^{-\frac{1}{2}\mathrm{i}kx_0}\\ \mathrm{e}^{-\frac{1}{2}\mathrm{i}kx_0} & \mathrm{e}^{\frac{1}{2}\mathrm{i}kx_0}\end{vmatrix}=0 \tag{1-68}$$

解式（1-68）可得

$$k=\frac{n\pi}{x_0}\quad（n\text{ 为整数}）\tag{1-69}$$

也就是，波数应该取由整数 n 限定的不连续值（$\lambda=2\pi/k$，λ 也同样）。由式（1-63）和式（1-64）给出的波数 k 与角频率 ω 的关系，可看出 ω 也不允许取任意值，而只能取不连续值。

由于把式（1-69）代入式（1-67）时，可得 $A_+=(-1)^{n+1}A_-$，所以 $\psi(x)$ 的一般解可写为

$$\psi=A_-\left[(-1)^{n+1}\mathrm{e}^{\mathrm{i}\frac{n\pi}{x_0}x}+\mathrm{e}^{-\frac{n\pi}{x_0}x}\right]\tag{1-70}$$

脚标 n 表示状态由 n 决定。

n 对应的角频率为 ω_n，由式（1-62）则有

$$\phi_n=\mathrm{e}^{-\mathrm{i}\omega_n t}\tag{1-71}$$

所以波函数 Ψ 为

$$\Psi_n=\psi_n\phi_n\tag{1-72}$$

1.2.4　波的叠加

1. 叠加原理

现在令 Ψ_1 和 Ψ_2 满足波动方程式(1-49)。如果 c_1、c_2 为常数，若 $\Psi=c_1\Psi_1+c_2\Psi_2$，很容易证明 Ψ 也满足式（1-49）的波动方程。再进一步，若 Ψ_1、Ψ_2、Ψ_2、…、Ψ_n 分别为波动方程的解时，那么它们的线性组合

$$\Psi=\sum_{i=1}^{n}c_i\Psi_i\tag{1-73}$$

也为波动方程式的解。式中 c_i 是常数。这就叫做叠加原理，是波动现象的特征之一。光和声波的干涉和衍射现象就是由于波的叠加产生的。

波函数绝对值的平方表示波的强度，如果利用式（1-73）则有

$$|\Psi|^2=|c_1\Psi_1|^2+|c_2\Psi_2|^2+c_1c_2\Psi_1\Psi_2+c_1c_2\Psi_1\Psi_2\tag{1-74}$$

合成波的强度不等于各自分波强度的简单相加，而是多出两个分波的函数的乘积项。此项就是干涉项。干涉项根据两个波函数的位相差，它的值从正到负有较大的变化。

2. 两简谐波叠加形成的波包

现在考察两个频率相近、振幅相同、同方向传播的简谐波的叠加。设两个波的波动表达式为

$$\Psi_1 = A\cos\left(\omega_1 t - k_1 x\right) \tag{1-75}$$

$$\Psi_2 = A\cos\left(\omega_2 t - k_2 x\right) \tag{1-76}$$

为了简单起见，假设它们在原点的初相位相等且为零（这里 A 即为振幅，与上面不同）。其合成波为

$$\Psi = \Psi_1 + \Psi_2 = 2A\cos\left(\frac{\omega_1 - \omega_2}{2}t - \frac{k_1 - k_2}{2}x\right)\cos\left(\frac{\omega_1 + \omega_2}{2}t - \frac{k_1 + k_2}{2}x\right) \tag{1-77}$$

合成的结果如图 1-18 所示。图中上图表示频率稍有不同的两列波，下图表示合成波（合成波振幅（虚线）随时间作周期性地变化）。由于 $\Delta\omega = \omega_1 - \omega_2 \ll \omega_1$ 或 ω_2，$\Delta k = k_1 - k_2 \ll k_1$ 或 k_2，所以 $\cos\left(\frac{\Delta\omega}{2}t - \frac{\Delta k}{2}x\right)$ 变化缓慢，如虚线所示的包络线；而 $\cos\left(\frac{\omega_1 + \omega_2}{2}t - \frac{k_1 + k_2}{2}x\right)$ 则表示一个小的波形。如果令 $\omega_m = \frac{\Delta\omega}{2}$，$k_m = \frac{\Delta k}{2}$，$\bar{\omega} = \frac{\omega_1 + \omega_2}{2}$，$\bar{k} = \frac{k_1 + k_2}{2}$，则式（1-77）可写为

$$\Psi = 2A\cos\left(\omega_m t - k_m x\right)\cos\left(\bar{\omega}t - \bar{k}x\right) = A_m(x,t)\cos\left(\bar{\omega}t - \bar{k}x\right) \tag{1-78}$$

可以把式（1-78）看成是一个角频率为 $\bar{\omega}$、波数为 $\bar{k}$ 的波，这个波的速度为

$$u_p = \frac{\bar{\omega}}{\bar{k}} \tag{1-79}$$

它的振幅 $A_m(x, t)$ 随时间和位置作比较缓慢的周期性变化。显然，这个变化的振幅 $A_m(x, t)$ 本身也具有沿 x 方向传播的简谐波的形式（如图 1-19 所示），它的角频率和波数分别为 ω_m 和 k_m，而传播速度为

$$u_g = \frac{\omega_m}{k_m} = \frac{\Delta\omega}{\Delta k} \tag{1-80}$$

通常称这样的振幅为调制的振幅。

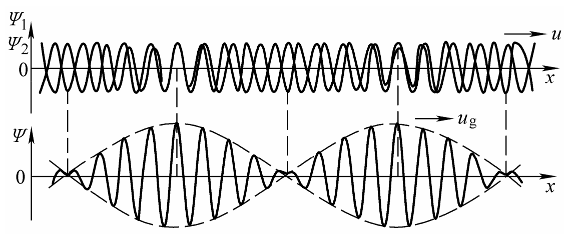


图 1-18　波包

这样，两个频率相近、等振幅的简谐波叠加的结果是一个振幅缓慢变化的波，它的角频率为 $\bar{\omega}$、波数为 $\bar{k}$，波速为 $u_p = \bar{\omega}/\bar{k}$。它的振幅的变化也像一个传播的

波，它的角频率为 $\Delta\omega/2$，波数为 $\Delta k/2$，波速 $u_g = \Delta\omega / \Delta k$。这样的合成波称为波包。波包有两个速度，$u_p$ 称为它的相速度，u_g 称为它的群速度（在图 1-19 中箭头以群速度 u_g 行进，白圈以相速度 u_p 行进）。群速度在研究光的散射以及微观粒子波动性等问题都是很重要的概念。

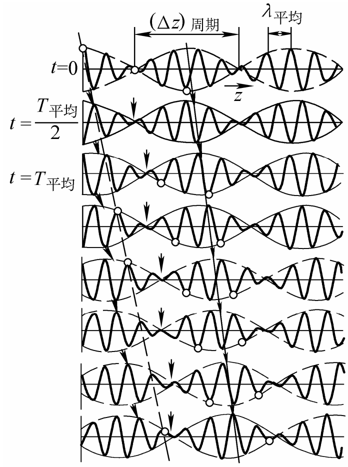


图 1-19　波包相速度和群速度

3. 两简谐波叠加形成的驻波

驻波是一种波形不向前传播的波，波形向前传播的波称之为行波。驻波是干涉的特例。在同一媒质中两列振幅相同的相干波（即振幅|A|、角频率 ω、波长 λ 均相同），在同一直线上沿相反方向传播时就叠加形成驻波。设有两列振幅相同的相干波分别沿 x 轴正方向和负方向传播，在两波中各选一个波峰，取两波相遇时 $t=0$，且 $x=0$，它们的表达式可写成

$$\Psi_1 = A\cos\left(\omega t - \frac{2\pi}{\lambda}x\right) \tag{1-81}$$

$$\Psi_2 = A\cos\left(\omega t + \frac{2\pi}{\lambda}x\right) \tag{1-82}$$

其合成波为

$$\Psi = \Psi_1 + \Psi_2 = 2A\cos\frac{2\pi}{\lambda}x\cos\omega t \tag{1-83}$$

式（1-83）中的空间坐标和时间坐标分开，是驻波的典型表达式。由式（1-83）可以看出，合成以后各点都在做同频率的简谐振动，但各点的振幅为$|2A\cos(x2\pi/\lambda)|$，即驻波的振幅与位置有关（与时间无关）。振幅最大发生在$|\cos(x2\pi/\lambda)| = 1$，即 $x2\pi/\lambda = k\pi$ 的各点，这些点称为波腹，它们的位置可由下式决定

$$x = k\frac{\lambda}{2}\quad (k = 0,\ \pm1,\ \pm2,\ \cdots) \tag{1-84}$$

亦即波腹在与原点距离为 $\lambda/2$ 的整数倍的位置处。同样，振幅的最小值发生在 $|\cos(x2\pi/\lambda)| = 0$，即 $x2\pi/\lambda = (2k+1)\pi/2$ 的各点，这些点称为波节，它们的位置可由下式决定

$$x = (2k+1)\frac{\lambda}{4}\quad (k = 0,\ \pm1,\ \pm2,\ \cdots) \tag{1-85}$$

亦即波节在与原点的距离为 $\lambda/4$ 的奇数倍的位置处。由式（1-84）和式（1-85）可算出相邻的两个波节和相邻的两个波腹之间的距离都是 $\lambda/2$。这一点为我们提供了一种测定波长的方法，只要测出相邻两波节或波腹之间的距离就可以确定原来两

列行波的波长。

式（1-83）中的振动因子为 $\cos\omega t$，似乎驻波中各点的振动的相都是相同的。其实不然，因为系数 $2A\cos(x2\pi/\lambda)$在 x 的值不同时是有正有负的。把相邻两个波节之间的各点叫做一段，则由余弦函数取值的规律可知，$\cos(x2\pi/\lambda)$的值对于同一段内的各点有相同的符号，对于分别在相邻两段内的两点则符号相反。以$|2A\cos(x2\pi/\lambda)|$作为振幅，这种符号的相同或相反就表明，在驻波中，同一段上的各点的振动同相，而相邻两段中的各点的振动反相。因此，驻波实际上就是分段振动现象。图 1-20 画出了驻波形成的物理过程，其中虚线表示向右传播的波 $\varPsi_1$ 的波形，点线表示向左传播的波 $\varPsi_2$ 的波形，点划线代表虚线和点线重合时的波形，实线是合成波 $\varPsi$ 的波形。由图 1-20 可见，在各个波节处，两波在任一时刻的位移恰好相消，在波腹处，两波的位移相长。在同一段上驻波各点振动同相，相邻两段中的各点振动反相。

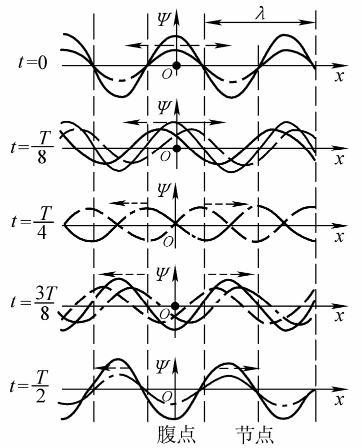


图 1-20　振幅相等、频率相同、传播方向相反的两列简谐波叠加而形成驻波

1.2.5　傅里叶解析

满足边界条件$\psi(x_0/2) = \psi(-x_0/2) =0$ 的波动方程式的一般解为

$$\varPsi(x,t) = \sum_{n=-\infty}^{\infty} c_n \mathrm{e}^{\mathrm{i}\left(\frac{n\pi}{x_0}x-\omega nt\right)} \tag{1-86}$$

当 $t=0$ 时的波函数 $\varPsi(x,0)$若是已知的，可求出系数 c_n。

$$\varPsi(x,0) = \sum_{n=-\infty}^{\infty} c_n \mathrm{e}^{\mathrm{i}\frac{n\pi}{x_0}x} \tag{1-87}$$

在式（1-87）的两边乘上$\mathrm{e}^{-\mathrm{i}\frac{m\pi}{x_0}x}$（$m$ 为整数），在$(-x_0, x_0)$区间积分，则成为

$$\int_{-x_0}^{x_0}\mathrm{e}^{-\mathrm{i}\frac{m\pi}{x_0}x}\Psi(x,0)\mathrm{d}x=\sum_{n=-\infty}^{\infty}c_n\int_{-x_0}^{x_0}\mathrm{e}^{\mathrm{i}\frac{(n-m)\pi}{x_0}x}\mathrm{d}x \tag{1-88}$$

式（1-88）右边的积分为

$$\left.\begin{aligned}&\int_{-x_0}^{x_0}\mathrm{e}^{\mathrm{i}\frac{(n-m)\pi}{x_0}x}\mathrm{d}x=\frac{-\mathrm{i}x_0}{(n-m)\pi}\left(\mathrm{e}^{\mathrm{i}(n-m)\pi}-\mathrm{e}^{-\mathrm{i}(n-m)\pi}\right)=0\qquad (n\neq m)\\&\int_{-x_0}^{x_0}\mathrm{e}^{\mathrm{i}\frac{(n-m)\pi}{x_0}x}\mathrm{d}x=\int_{-x_0}^{x_0}\mathrm{d}x=2x_0\qquad (n=m)\end{aligned}\right\} \tag{1-89}$$

所以由式（1-88）得

$$c_n=\frac{1}{2x_0}\int_{-x_0}^{x_0}\mathrm{e}^{-\mathrm{i}\frac{n\pi}{x_0}x}\Psi(x,0)\,\mathrm{d}x \tag{1-90}$$

因而，可得到任意时间的波函数为

$$\Psi(x,t)=\frac{1}{2x_0}\sum_{n=-\infty}^{\infty}\left\{\int_{-x_0}^{x_0}\mathrm{e}^{-\mathrm{i}\frac{n\pi}{x_0}x}\Psi(x,0)\mathrm{d}x\right\}\mathrm{e}^{\mathrm{i}\left(\frac{n\pi}{x_0}x-\omega nt\right)} \tag{1-91}$$

根据傅里叶级数展开的波函数就是式（1-86）。展开系数 c_n 可由式（1-90）具体计算。

如上所述，因为函数$\mathrm{e}^{\mathrm{i}\frac{n\pi}{x_0}x}$在（$-x_0, x_0$）区间中满足式（1-89）那样的关系，所以能够进行傅里叶展开，并能求出它的系数。这个关系称为正交关系。即函数$\mathrm{e}^{\mathrm{i}\frac{n\pi}{x_0}x}$（$n$ 为整数）在（$-x_0, x_0$）区间中，建立了正交关系。如果区间任意改变时不满足正交关系，也就不能求出系数 c_n。

如果体系包含一切相互正交的函数时，一般认为体系是完全的，叫做完全正交体系。任意函数，可以用完全正交体系函数的线性组合表示。

其次，考虑决定区间大小的 x_0 无限大的情况。这时由于$\Delta k = \pi/x_0$变化成无限小，随着 n 的变化，k 值大体上是连续变化。用Δk，式（1-90）就可写成

$$c_n=\frac{\Delta k}{2\pi}\int_{-x_0}^{x_0}\mathrm{e}^{-\mathrm{i}k_n x}\Psi(x,0)\mathrm{d}x \tag{1-92}$$

如定义如下的函数

$$c(k)=\sqrt{2\pi}\lim_{x_0\to\infty}\frac{c_n}{\Delta k} \tag{1-93}$$

则得

$$c(k)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\mathrm{e}^{-\mathrm{i}kx}\Psi(x,0)\mathrm{d}x \tag{1-94}$$

如果运用 $c(k)$，则式（1-87）求 $x_0\to\infty$的极限。

$$\Psi(x,0)=\lim_{x_0\to\infty}\sum_n c_n\mathrm{e}^{\mathrm{i}\frac{n\pi}{x_0}x}=\lim_{\Delta k\to 0}\sum_n\frac{c_n}{\Delta k}\mathrm{e}^{\mathrm{i}k_n x}\Delta k=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}c(k)\,\mathrm{e}^{\mathrm{i}kx}\mathrm{d}k \tag{1-95}$$

因而可得

$$\Psi(x,t)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}c(k)\mathrm{e}^{\mathrm{i}(kx-\omega t)}\mathrm{d}k \tag{1-96}$$

可见式（1-94）和式（1-95）互为傅里叶变换关系。$c(k)$表示波数在 k 和 k+dk 之间的波以任意一种比例存在的情况，称为分布函数。

如果把 $c(k)\mathrm{e}^{-\mathrm{i}\omega t}$ 重新写成 $c(k, t)$，则式（1-94）、式（1-95）可表示为如下的对称形式

$$c(k,t)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\mathrm{e}^{-\mathrm{i}kx}\Psi(x,t)\mathrm{d}k \tag{1-97}$$

$$\Psi(x,t)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}kx}c(k,t)\mathrm{d}k \tag{1-98}$$

波在三维空间传播时，设波矢为 $\boldsymbol{k}(k_x, k_y, k_z)$，位矢为 $\boldsymbol{r}(x, y, z)$，式（1-98）可改写为

$$c(\boldsymbol{k},t)=\frac{1}{\sqrt{8\pi^3}}\iiint_{-\infty}^{\infty}\mathrm{e}^{-\mathrm{i}\boldsymbol{kr}}\Psi(\boldsymbol{r},t)\mathrm{d}x\mathrm{d}y\mathrm{d}z \tag{1-99}$$

$$\Psi(\boldsymbol{r},t)=\frac{1}{\sqrt{8\pi^3}}\iiint_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}\boldsymbol{kr}}c(\boldsymbol{k},t)\,\mathrm{d}k_x\mathrm{d}k_y\mathrm{d}k_z \tag{1-100}$$

1.2.6　波束（波包）

现在讨论在空间的有限部分中波函数不为零的情况。高斯函数（图 1-21a）如下

$$\Psi\left(x,0\right)=\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2(\Delta x)^2}} \tag{1-101}$$

这里 Δx 为波包宽度。如果把式（1-101）代入式（1-94），则得

$$c(k)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{e}^{-\mathrm{i}kx-\frac{x^2}{2(\Delta x)^2}}\mathrm{d}x=\frac{\mathrm{e}^{-\frac{(\Delta x)^2k^2}{2}}}{2\pi}\int_{-\infty}^{\infty}\mathrm{e}^{-\frac{z^2}{2(\Delta x)^2}}\mathrm{d}z=\frac{\Delta x}{\sqrt{2\pi}}\mathrm{e}^{-\frac{(\Delta x)^2k^2}{2}} \tag{1-102}$$

$c(k)$和 $\Psi(x,0)$一样是高斯函数，如图 1-21 所示。但和 $\Psi(x,0)$函数不同，$c(k)$的扩展由 $\frac{1}{\Delta x}$ 得到。一般情况下，当 $\Psi(x, 0)$的扩展越大，$c(k)$的扩展就变小了，反之，如果 $\Psi(x,0)$越局部化则 $c(k)$的扩展就越大。$\Psi(x,0)$函数等于 $1/e$ 的 x 值为 $\sqrt{2}\Delta x$，但如果 $c(k)$等于 $1/e$ 的 k 值为 $\sqrt{2}\Delta k$ 时，从式（1-102）的关系得

$$\Delta x \Delta k = 1 \tag{1-103}$$

该式表示波的空间扩展和波数的扩展之间存在着互补关系。为了使波定域，波数应该有某种程度的分布，把这样的波叫做波束。换句话说，所谓波束就是合成各种成分波而得到的。物质波的测不准原理可以认为是把波束大小和波数扩展的一般互补关系适用于物质波的原理。

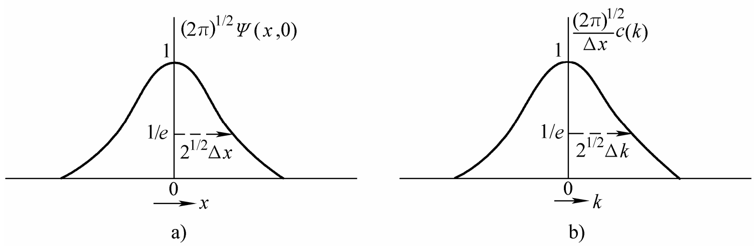


图 1-21　波束（高斯函数）

a）$\Psi(x, 0)$函数　b）$c(k)$函数

1.2.7　波束的运动

1.2.6 节对时间为一定的波束进行了讨论，下面再讨论一下波束随时间是怎样变化的。波束在一般的情况下以式（1-86）或式（1-96）的形式表示，在子波的相速相同时，波束不改变它的传播形式。即在式（1-86）中有如下关系式

$$\frac{\omega_1}{k_1} = \frac{\omega_2}{k_2} = \cdots = \frac{\omega_n}{k_n} = u_p \tag{1-104}$$

则变成

$$\Psi(x,t) = \sum_{n=-\infty}^{\infty} c_n \mathrm{e}^{\mathrm{i}k_n(x-u_p t)} = F(x-u_p t) \tag{1-105}$$

如果在式（1-96）中代入 $\omega / k = u_p$ 时，则得

$$\Psi(x,t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} c(x) \mathrm{e}^{\mathrm{i}k(x-u_p t)} \mathrm{d}k = G(x-u_p t) \tag{1-106}$$

F、G 是 $x-u_p t$ 的函数，波束以 u_p 的速度传播。ω/k 为一定时，叫做非色散性。如 ω/k 不是常数时，$\Psi(x, t)$不能作为 $x-u_p t$ 的函数表现出来，波束随着时间变化边改变形式边移动，把这样的情况叫做色散性。决定波束传播的形式是 $\omega(k)$的函数形

式。$\omega(k)$是 k 的一次函数时为非色散性，高次函数时成为色散性。在这种意义上说 $\omega(k)-k$ 的曲线叫做色散曲线。

在色散性介质中，因为波束随时间变化而改变它的形状，所以严格意义上来说，不能定义波束的移动速度，但在极短时间内，则可以类似式（1-80）定义 $u_g = \mathrm{d}\omega/\mathrm{d}k$ 为群速度，认为是波束的移动速度。在非色散性时，$u_p = u_g$ 成立。以波失 $\boldsymbol{k}$ 给出时，群速度可表示为

$$u_g = \nabla_{\boldsymbol{k}}\omega(\boldsymbol{k}) \tag{1-107}$$

∇_k 是(k_x, k_y, k_z)的微分算符。

1.3　统计力学概要

通常观测到的宏观性质都是大量粒子运动情况的平均结果。正是统计力学把各个粒子的行为与大量粒子的集合的宏观性质联系了起来。本节将概要地介绍后面论述所必须的统计力学知识。

在大量粒子的集合中，包含着具有不同能量和速度的粒子。设在能量 ε 和 $\varepsilon+\mathrm{d}\varepsilon$ 之间的变化区域 $\mathrm{d}\varepsilon$ 中的粒子数为 $n(\varepsilon)\mathrm{d}\varepsilon$，总粒子数为 N，将 $a(\varepsilon)$称为能量分布函数。$a(\varepsilon)$为

$$a(\varepsilon) = \frac{n(\varepsilon)}{N} \tag{1-108}$$

$a(\varepsilon)$表示在 ε 附近单位能量变化区域中的粒子在总粒子中所占的比例。根据定义有 $N = \int_0^\infty n(\varepsilon)\mathrm{d}\varepsilon$，因此有

$$\int_0^\infty a(\varepsilon)\mathrm{d}\varepsilon = 1 \tag{1-109}$$

即分布函数是归一化的。设在能量 ε 到 $\varepsilon+\mathrm{d}\varepsilon$ 之间的粒子的能级数为 $\rho(\varepsilon)\,\mathrm{d}\varepsilon$，则称 $\rho(\varepsilon)$为能量状态密度，或简称状态密度。另外，能量为 ε 的能级被粒子占据的几率 $f(\varepsilon)$称为占有率。因为在能量 ε 到 $\varepsilon+\mathrm{d}\varepsilon$ 之间的粒子数 $n(\varepsilon)\mathrm{d}\varepsilon$ 等于能量 ε 到 $\varepsilon+\mathrm{d}\varepsilon$ 之间的能级数和占有率 $f(\varepsilon)$的乘积，所以分布函数可按下式由状态密度与占有几率的乘积给出

$$a(\varepsilon) = \frac{n(\varepsilon)}{N} = \frac{1}{N}\rho(\varepsilon)f(\varepsilon) \tag{1-110}$$

根据统计的方法，占有率 $f(\varepsilon)$可由玻耳兹曼分布、玻色分布和费米分布三种方法求得。

1.3.1　系统的微观运动状态描述

所谓系统的微观运动状态就是它的力学运动状态。这里限于讨论由全同和近

独立粒子组成的系统。全同粒子组成的系统就是由具有完全相同的属性（相同的质量、电荷、自旋等等）的同类粒子组成的系统。近独立粒子组成的系统，是指系统中粒子之间相互作用很弱，相互作用的平均能量远小于单个粒子的平均能量，因而可以忽略粒子之间的相互作用，整个系统的能量为单个粒子的能量之和。理想气体就是由近独立粒子组成的系统。理想气体的分子，除了相互碰撞的瞬间，都可以认为没有相互作用。

1. 经典力学系统的微观运动状态描述和玻耳兹曼系统

设粒子的自由度为 r。在任一时刻，第 i 个粒子的力学运动状态由 r 个广义坐标 q_{i1}，q_{i2}，q_{i3}，…，q_{ir} 和 r 个广义动量 p_{i1}，p_{i2}，p_{i3}，…，p_{ir} 的数值确定。当组成系统的 N 个粒子在某一时刻的力学运动状态都确定时，整个系统在该时刻的微观运动状态也就确定了。因此确定系统的微观运动状态需要 $2Nr$ 个变量，这 $2Nr$ 个变量就是 q_{i1}，q_{i2}，q_{i3}，…，q_{ir} 和 p_{i1}，p_{i2}，p_{i3}，…，p_{ir}($i = 1$，2，…，N)。在经典物理中，全同粒子是可以分辨的。这是因为，经典粒子的运动是轨道运动，原则上是可以被跟踪的。只要确定每一粒子在初始时刻的位置，原则上就可以确定每一粒子在其后任一时刻的位置。所以尽管全同粒子的属性完全相同，原则上仍然可以辨认。既然全同粒子可以分辨，如果在含有多个全同粒子的系统中，将两个粒子的运动状态加以交换，例如第 i 个粒子和第 j 个粒子的运动状态本来分别是(q_1'，q_2'，q_3'，…，q_r'；p_1'，p_2'，p_3'，…，p_r')和(q_1''，q_2''，q_3''，…，q_r''；p_1''，p_2''，p_3''，…，p_r'')，如果将它们的运动状态加以交换，使第 i 个粒子的运动状态为(q_1''，q_2''，q_3''，…，q_r''；p_1''，p_2''，p_3''，…，p_r'')，第 j 个粒子的运动状态为(q_1'，q_2'，q_3'，…，q_r'；p_1'，p_2'，p_3'，…，p_r')。在交换前后，系统的力学运动状态是不同的。

一个粒子在某一时刻的力学运动状态可用 μ 空间中的一个点表示。由 N 个全同粒子组成的系统在某一时刻的微观运动状态可在 μ 空间中用 N 个点表示。根据前面的讨论可知，如果交换两个代表点在 μ 空间的位置，相应的系统的微观状态是不同的。在统计物理学发展的早期（远在量子力学建立以前），玻耳兹曼把粒子看作是可以分辨的，并导出了这种粒子的统计分布。现在我们把由可分辨的全同近独立粒子组成，且处在一个个体量子态上的粒子数不受限制的系统称作玻耳兹曼系统。

2. 量子系统微观运动状态的描述和玻色系统、费米系统

微观粒子全同性原理指出，全同粒子是不可分辨的，在含有多个全同粒子的系统中，将任何两个全同粒子加以对换，不改变整个系统的微观运动状态。这原理与经典物理关于全同粒子可以分辨的论断是完全不同的。导致完全不同的论断的根本原因是，经典粒子的运动是轨道运动，原则上可以跟踪经典粒子的运动而加以辨认，而量子粒子具有波粒二象性，它的运动不是轨道运动，原则上不可能跟踪量子粒子的运动。假设在 t =0 时确知两个粒子的位置，由于与这两个粒子相

联系的波动迅速扩散而互相重叠，在 $t>0$ 时在某一地点发现粒子时，已经不能辨认到底是第一个还是第二个粒子了。图 1-22a、b 示意地表示两个粒子遵从经典力学和量子力学的区别。

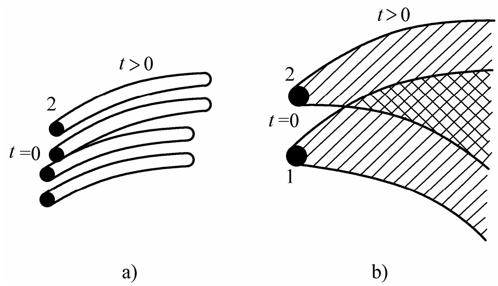


图 1-22　全同粒子

a）经典力学的情景　b）量子力学的情景

假如全同粒子可以分辨，确定由全同近独立粒子组成的系统的微观运动状态归结为确定每一个粒子的个体量子态。对于不可分辨的全同粒子，确定由全同近独立粒子组成的系统的微观状态归结为确定每一个体量子态上的粒子数。例如，确定 He 气的微观状态，归结为确定由每一组量子数 n_x、n_y、n_z 所表征的个体量子态上各有多少个 He 原子。

在讨论量子粒子怎样占据各个个体量子态时，还有一个原则问题必须给予考虑。自然界中微观粒子可分为两类，称为玻色子和费米子。在“基本”粒子中，自旋量子数为半整数的，例如电子、质子、中子等自旋量子数都是 1/2，是费米子；自旋量子数是整数的，例如光子自旋量子数为 1，π 介子自旋量子数为零，是玻色子。在原子核、原子和分子等复合粒子中，凡是由玻色子构成的复合粒子是玻色子，由偶数个费米子构成的复合粒子也是玻色子，由奇数个费米子构成的复合粒子是费米子。

由费米子组成的系统称为费米系统，遵从泡利不相容原理。泡利不相容原理说，在含有多个全同近独立的费米子的系统中，一个个体量子态最多能容纳一个费米子。由玻色子组成的系统称为玻色系统，不受泡利不相容原理的约束。这就是说，由多个全同近独立的玻色子组成的玻色系统中，处在同一个体量子态的玻色子数目是不受限制的。

3. 玻耳兹曼系统、玻色系统和费米系统的区别

现在举一个简单的例子说明玻耳兹曼系统、玻色系统和费米系统的区别。设系统含有两个粒子，粒子的个体量子态有 3 个。现在考察，对于玻耳兹曼系统、玻色系统和费米系统各有哪些可能的微观状态。

玻耳兹曼系统，粒子可以分辨，每一个体量子态能够容纳的粒子数不受限制。

以 A、B 表示可以分辨的两个粒子，它们占据 3 个个体量子态可以有以下的方式

量子态 1	量子态 2	量子态 3
A B		
	A B	
		A B
A	B	
B	A	
	A	B
	B	A
A		B
B		A

因此，对于玻耳兹曼系统，可以有 9 个不同的状态。

玻色系统，粒子不可分辨，每一个个体量子态所能容纳的粒子数不受限制，由于粒子不可分辨，令 A = B。两个粒子占据 3 个个体量子态有以下的方式

量子态 1	量子态 2	量子态 3
A A		
	A A	
		A A
A	A	
	A	A
A		A

因此，对于玻色系统，可以有 6 个不同的状态。

费米系统，粒子不可分辨，每一个个体量子态最多能容纳一个粒子。两个粒子占据 3 个个体量子态有以下的方式

量子态 1	量子态 2	量子态 3
A	A	
	A	A
A		A

因此，对于费米系统，可以有 3 个不同的微观状态。

在经典力学基础上建立的统计物理学称为经典统计物理学，在量子力学基础上建立的统计物理学称为量子统计物理学。两者在统计原理上是相同的，区别在于对微观运动状态的描述。微观粒子实际上遵从量子力学的运动规律。不过在一定的极限条件下，经典统计在一定条件下还是有意义的。

1.3.2 分布和微观状态

只要知道各个微观状态出现的概率，就可以用统计方法求微观量的统计平均

值。因此确定各微观状态出现的概率是统计物理的根本问题。对于这个问题，玻耳兹曼在 19 世纪 70 年代提出了著名的等概率原理。等概率原理认为，对于处在平衡状态的孤立系统，系统各个可能的微观状态出现的概率是相等的。等概率原理在统计物理中是一个基本假设。它的正确性由它的种种推论都与客观实际相符而得到肯定。

设有一个系统，由大量全同近独立的粒子组成，具有确定的粒子数 N、能量 E 和体积 V。以 ε_l 表示粒子的能级（$l=1$，2，⋯），ω_l 表示能级 ε_l 的简并度（即量子状态数），a_l 表示 N 个粒子中在 ε_l 能级的粒子数，即

能级	ε_1,	ε_2,	ε_3,	⋯,	ε_l,	⋯
简并度	ω_1,	ω_2,	ω_3	⋯,	ω_l,	⋯
粒子数	a_1	a_2	a_3	⋯,	a_l,	⋯

为了书写方便，以符号 $\{a_l\}$ 表示数列 a_1, a_2, a_3, ⋯, a_l, ⋯，称为一个分布。显然，对于具有确定的 N、E、V 的系统，分布必须满足条件

$$\sum_l a_l = N\ ,\quad \sum_l a_l\varepsilon_l = E \tag{1-111}$$

分布和微观状态是两个不同的概念。给定一个分布 $\{a_l\}$，只确定了在每一个能级 ε_l 上的粒子数 a_l。如前所述，对于玻色系统和费米系统，确定系统的微观状态要求确定处在每一个个体量子态上的粒子数。因此在分布给定后，要确定玻色（费米）系统的微观状态，还必须对每一个能级确定 a_l 个粒子占据其 ω_l 个量子态的方式。对于玻耳兹曼系统，确定系统的微观状态要求确定每一个粒子的个体量子态。因此在分布给定后，为了确定玻耳兹曼系统的微观状态，还必须确定处在各能级 ε_l 上的是哪个 a_l 粒子，以及在每一能级 ε_l 上 a_l 个粒子占据其 ω_l 个量子态的方式。由此可见，与一个分布 $\{a_l\}$ 相应的系统的微观状态往往是很多的。这微观状态数对于玻耳兹曼系统、玻色系统和费米系统显然不同，下面分别加以讨论。

1. 玻耳兹曼系统

对于玻耳兹曼系统，粒子可以分辨，可以对粒子加以编号。a_l 个编了号的粒子占据能级 ε_l 上的 ω_l 个量子态时，第一个粒子可以占据 ω_l 个量子态中的任何一态，有 ω_l 种可能的占据方式。由于一个量子态可以容纳的粒子数不受限制，在第一个粒子占据了某一个量子态以后，第二、第三、⋯个粒子仍然有 ω_l 种可能的占据方式。这样，a_l 个编了号的粒子占据 ω_l 个量子态共有 ${\omega_l}^{a_l}$ 种可能的占据方式。这就是说，在玻耳兹曼系统中，a_l 个粒子占据能级 ε_l 上的 ω_l 个量子态时，是彼此独立、互不关联。a_1, a_2, a_3, ⋯, a_l, ⋯个编了号的粒子分别占据能级 $\varepsilon_1, \varepsilon_2, \varepsilon_3$, ⋯, ε_l, ⋯上的各量子态共有 $\prod_l {\omega_l}^{a_l}$ 种方式。玻耳兹曼系统的粒子既然可以分辨，交换粒子将

给出系统的不同状态。将 N 个粒子加以交换，不管是否在同一能级上，交换数是 $N!$。在这交换数中应除去在同一能级上 a_l 个粒子的交换数 $a_l!$，因此得因子 $N!/\prod_l a_l!$。所以，对于玻耳兹曼系统，与分布 $\{a_l\}$ 相应的系统的微观状态数是

$$\Omega_{\mathrm{M.B.}}=\frac{N!}{\prod_l a_l!}\prod_l \omega_l^{a_l} \tag{1-112}$$

2. 玻色系统和费米系统

对于玻色系统，粒子不可分辨，每一个个体量子态能够容纳的粒子数不受限制。首先计算 a_l 个粒子占据能级 ε_l 上的 ω_l 个量子态有多少种可能的方式。为了计算这个数目，量子态 1, 2, 3, …分别用①, ②, ③, …表示，用⊙表示全同粒子，将它们排成一行，使最左方为量子态 1。图 1-23 表示 5 个量子态和 10 个粒子的一种排列，这样的排列表示在量子态 1 上有 2 个粒子，在量子态 2 上有 1 个粒子，在量子态 3 上没有粒子，在量子态 4 上有 3 个粒子，在量子态 5 上有 4 个粒子。由于最左方固定为量子态 1，其余的量子态和粒子的总数是$(\omega_l + a_l-1)$个，将它们加以排列共有$(\omega_l + a_l -1)!$种方式（$\omega_l = 5$，$a_l = 10$，14！种）。因为粒子是不可分辨的，应除去粒子之间的相互交换数 $a_l!$（$\omega_l =5$，$a_l = 10$，10！种）和量子态之间的相互交换数$(\omega_l -1)!$（$\omega_l = 5$，$a_l = 10$，4！种）。这样便可得到，a_l 个粒子占据能级 ε_l 上的 ω_l 个量子态，有$(\omega_l+a_l -1)!/[a_l!(\omega_l -1)!]$种可能的方式。将各能级的结果相乘，就得到玻色系统与分布 $\{a_l\}$ 相应的微观状态数为

①⊙⊙②⊙③④⊙⊙⊙⑤⊙⊙⊙⊙

图 1-23　量子态和粒子的排列

$$\Omega_{\mathrm{B.E.}}=\prod_l \frac{(\omega_l + a_l -1)!}{a_l!(\omega_l -1)!} \tag{1-113}$$

对于费米系统，粒子不可分辨，每一个个体量子态最多只能容纳一个粒子。a_l 个粒子占据能级 ε_l 上的 ω_l 个量子态，相当于从 ω_l 个量子态中挑出 a_l 个来为粒子所占据（注意 $\omega_l \geqslant a_l$），有 $\omega_l!/[a_l!(\omega_l - a_l)!]$种可能的方式。将各能级的结果相乘，就得到费米系统与分布 $\{a_l\}$ 相应的微观状态数为

$$\Omega_{\mathrm{F.D.}}=\prod_l \frac{\omega_l!}{a_l!(\omega_l - a_l)!} \tag{1-114}$$

1.3.3　玻尔兹曼分布

在求解玻尔兹曼分布、玻色分布和费米分布前，先说一下最概然分布和斯特令公式。温度、压力、平均能量等相同的一个宏观状态，从微观角度看，对应着各种粒子的多种大量的组合方式（位置、动量等的组合）。与给定的宏观状态对应

的不同的微观状态数 Ω，称为该宏观状态的热力学概率。玻尔兹曼提出，在热力学概率和熵 S 之间存在如下的关系

$$S = k\ln\Omega \tag{1-115}$$

称为玻尔兹曼关系，k 为玻尔兹曼常数。在热力学中，熵最大的状态是稳定的状态，所以 Ω 最大的状态给出稳定的热平衡状态。平衡态是出现概率最大，即最概然的宏观态。平衡态对应的分布是最概然分布。在数学上，有一近似公式，即当 m 是一个远大于 1 的整数时，有

$$\ln m! \approx m(\ln m - 1) \tag{1-116}$$

即斯特令公式。

对玻尔兹曼系统中的微观状态（热力学概率）公式（1-112）两边取自然对数，得

$$\ln\Omega_{\mathrm{M.B.}} = \ln N! - \sum_l \ln a_l! + \sum_l a_l \ln\omega_l \tag{1-117}$$

假设所有的 a_l 都很大（当然 N 也很大），根据斯特令公式（1-116），式（1-117）可化为（将 $\Omega_{\mathrm{M.B.}}$ 简记为 Ω）

$$\begin{aligned}\ln\Omega &= N(\ln N - 1) - \sum_l a_l(\ln a_l - 1) + \sum_l a_l\ln\omega_l \\ &= N\ln N - \sum_l a_l\ln a_l + \sum_l a_l\ln\omega_l\end{aligned} \tag{1-118}$$

求熵最大，即求 $\ln\Omega$ 最大。为求得使 $\ln\Omega$ 为极大的分布，令各 a_l 有 δa_l 的变化，$\ln\Omega$ 将因而有 $\delta\ln\Omega$ 的变化。使 $\ln\Omega$ 为极大的分布，$\{a_l\}$ 必使 $\delta\ln\Omega = 0$，即

$$\delta\ln\Omega = -\sum_l(\ln a_l + 1)\delta a_l + \sum_l \ln\omega_l\delta a_l = -\sum_l \ln\left(\frac{a_l}{\omega_l}\right)\delta a_l = 0 \tag{1-119}$$

这些 δa_l 不完全是独立的，它们必须满足条件

$$\delta N = \sum_l \delta a_l = 0\text{，}\quad \delta E = \sum_l \varepsilon_l\delta a_l = 0 \tag{1-120}$$

用拉格朗日待定乘子 α 和 β 分别乘式（1-120）中的两个式子，并由式（1-119）可得

$$\delta\ln\Omega - \alpha\delta N - \beta\,\delta E = -\sum_l \ln\left(\frac{a_l}{\omega_l} + \alpha + \beta\varepsilon_l\right)\delta a_l = 0 \tag{1-121}$$

根据拉格朗日乘子法原理，每个 δa_l 的系数都等于零，所以得

$$\ln\frac{a_l}{\omega_l} + \alpha + \beta\varepsilon_l = 0 \tag{1-122}$$

即

$$a_l = \omega_l \mathrm{e}^{-\alpha-\beta\varepsilon_l} \quad （1\text{-}123）$$

式（1-123）给出玻尔兹曼系统中粒子的最概然分布，称为麦克斯韦—玻尔兹曼分布或玻尔兹曼分布。拉格朗日待定乘子 α 和 β 由式（1-111）确定，即

$$N = \sum_l \omega_l \mathrm{e}^{-\alpha-\beta\varepsilon_l}, \quad E = \sum_l \varepsilon_l \omega_l \mathrm{e}^{-\alpha-\beta\varepsilon_l} \quad （1\text{-}124）$$

式（1-123）给出在最概然分布下处在能级 ε_l 的粒子数。能级 ε_l 有 ω_l 个量子态，处在其中任何一个量子态的平均粒子数应该是相同的。因为 a_l/ω_l 是能量为 ε_l 的一个量子状态被粒子填充的几率，所以它就是占有率 $f(\varepsilon_l)$。另外 β 还可以热力学定义求得，有热力学定义可求得 $\beta = 1/(kT)$，k 即玻尔兹曼常数。因而式（1-123）可写成一般的表达式

$$f(\varepsilon) = A\mathrm{e}^{-\frac{\varepsilon}{kT}} \quad （1\text{-}125）$$

这里 $A=\mathrm{e}^{-\alpha}$

1.3.4 玻色分布和费米分布

对玻色系统中的微观状态（热力学概率）公式（1-113）两边取自然对数，得

$$\ln \Omega_{\mathrm{B.E.}} = \sum_l [\ln(\omega_l + a_l - 1)! - \ln a_l! - \ln(\omega_l - 1)!] \quad （1\text{-}126）$$

假设 $a_l \gg 1$，$\omega_l \gg 1$，因而 $a_l + \omega_l - 1 \approx a_l + \omega_l$，$\omega_l - 1 \approx \omega_l$，并根据斯特令公式（1-116），式（1-126）可化为（将 $\Omega_{\mathrm{B.E.}}$简记为 Ω）

$$\ln \Omega = \sum_l [(\omega_l + a_l)\ln(\omega_l + a_l) - a_l \ln a_l - \omega_l \ln \omega_l] \quad （1\text{-}127）$$

类似玻尔兹曼分布的求法，可求得

$$a_l = \frac{\omega_l}{\mathrm{e}^{\alpha+\beta\varepsilon_l} - 1} \quad （1\text{-}128）$$

式（1-128）给出玻色系统中粒子的最概然分布，称为玻色—爱因斯坦分布或玻色分布。

对费米系统中的微观状态（热力学概率）公式（1-114）两边取自然对数，得

$$\ln \Omega_{\mathrm{F.D.}} = \sum_l [\ln \omega_l! - \ln a_l! - \ln(\omega_l - a_l)!] \quad （1\text{-}129）$$

假设 $a_l \gg 1$，$\omega_l \gg 1$，$\omega_l - a_l \gg 1$，式（1-129）可近似为（将 $\Omega_{\mathrm{B.E.}}$简记为 Ω）

$$\ln \Omega = \sum_l [\omega_l \ln \omega_l - a_l \ln a_l - (\omega_l - a_l)\ln(\omega_l - a_l)] \quad （1\text{-}130）$$

可求得

$$a_l = \frac{\omega_l}{e^{\alpha+\beta\varepsilon_l}+1} \tag{1-131}$$

式（1-131）给出费米系统中粒子的最概然分布，称为费米—狄拉克分布或费米分布。

式（1-128）和式（1-131）分别给出玻色系统和费米系统在最概然分布下处在能级 ε_l 的粒子数。能级 ε_l 有 ω_l 个量子态，处在其中任何一个量子态的平均粒子数应该是相同的。因为 a_l/ω_l 是能量为 ε_l 的一个量子状态被粒子填充的几率，所以它就是占有率 $f(\varepsilon_l)$。另外 β 还可以热力学定义求得，由热力学定义可求得 $\beta = 1/(kT)$，并设 $\alpha = -\zeta/(kT)$，k 即玻尔兹曼常数。则有

$$f(\varepsilon) = \frac{1}{e^{(\varepsilon-\zeta)/kT}\pm 1} \tag{1-132}$$

正号为费米—狄拉克分布，负号为玻色—爱因斯坦分布。

图 1-24 示出了费米—狄拉克统计的分布函数，系统常数 ζ 称为费米能级。在 $T\to 0$ 的低温下，能量小于 ζ 的所有能级都被粒子填满，$f(\varepsilon) = 1$，能量大于 ζ 的所有能级都完全没有粒子，$f(\varepsilon) = 0$。随着温度的升高，如图 1-24 所示，粒子分布的界线逐渐模糊起来。

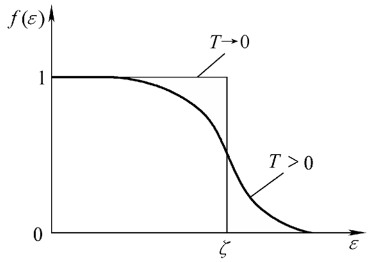


图 1-24　费米—狄拉克分布

无论是玻色—爱因斯坦分布还是费米—狄拉克分布，在能量很高（$\varepsilon-\zeta \gg kT$）时，粒子的分布都近似等于麦克斯韦—玻尔兹曼分布。

在粒子可以分辨的假设下导出玻尔兹曼分布。自然界中有些系统可以看作由定域的粒子组成，例如晶体中的原子或离子定域在其平衡位置附近作微振动，这些粒子虽然就其量子本性来说是不可分辨的，但可以根据其位置而加以区分。在这意义下可以将定域粒子看作可以分辨的粒子。因此由定域粒子组成的系统（称为定域系统）遵从玻尔兹曼分布。定域系统和满足经典极限条件的玻色（费米）系统虽然遵从同样的分布，但它们的微观状态数是不同的。

第 2 章　原子结构、化学键及晶体结构

自然界中为什么会形成如此繁多的物质，而它们又具有各种特性与功能，根源在于物质内部的组成及其结构。因此必须从更深的层次，从微观的角度来研究物质。为此应首先了解原子的结构，特别是核外电子的运动规律，继而了解物质内部的原子和分子等是如何通过化学键等相互作用力结合在一起的，从而为设计和合成具有特定结构性能的新化合物及新材料指明途径。

2.1　原子结构

原子是由带正电荷的原子核和绕原子核运动的带负电荷的电子所构成的。原子核的大小为 10^{-14}m 的数量级，与一般分子或晶体中的原子间距（约为 10^{-10}m 数量级）相比是很小的；与原子的核外电子所占的范围（约为 10^{-9}～10^{-10}m 数量级）相比也是很小的。原子核中又包含有质子及中子，质子与中子具有几乎相同的质量，即相当于氢原子的质量，或相当于电子质量（9.1085×10^{-28}g）的 1840 倍。因此原子核虽小，但却几乎集中了原子全部的质量。质子具有正电荷，每个质子所带电荷的数值与一个电子所带电荷的数值相等，但符号相反。一个原子中的质子数往往是与核外电子数相等的，这就使原子保持为中性。如果原子中质子数多于电子数，则为正离子，反之，则为负离子。

在普通化学反应中，一般都认为原子核是不可分的，并不涉及原子核的变化，只是核外电子的运动状态和分布发生变化。核外电子的运动状态与物质性质有着密切的关系，这是下面要着重讨论的问题。

2.1.1　氢原子

在本节中将讨论最简单的原子——氢原子的问题。它的结果可以很方便地推广到类氢原子（如氘、氚、一次电离的氦 He^{+}、二次电离的锂 Li^{++}等等）的情况，它是了解复杂原子及分子结构的基础。

氢原子包含一个原子核和一个电子，由于原子核的质量比电子的质量大很多，可以近似地把原子核看成是静止的。又因为要研究的只是原子的内部状态，而不是它的整体运动，所以，只研究电子相对于原子核的运动就可以了。

1. 氢原子的定态薛定谔方程

氢原子中的电子在原子核的库仑场中运动。原子核的电量为$+e$，电子的电量

为$-e$，所以电子的势能为

$$U(r) = -\frac{e^2}{4\pi\varepsilon_0}\frac{1}{r} = -\frac{{e_s}^2}{r} \tag{2-1}$$

式中，$e_s = e\big/\sqrt{4\pi\varepsilon_0}$；$r$ 为电子距原子核的距离；ε_0 为真空介电常数。那么电子的（亦即氢原子）定态薛定谔方程为

$$\frac{\hbar^2}{2\mu}\left\{\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right\}\psi + \left(\varepsilon + \frac{{e_s}^2}{r}\right)\psi = 0 \tag{2-2}$$

式中 μ 为电子与原子核的折合质量，$\mu = Mm_e/(M+m_e)$，M 为原子核质量，m_e 为电子质量。但由于原子核的质量比电子的质量大得多，因而，μ 即为电子的质量 m_e。ε 为电子的总能量（势能与动能之和）。由于 $U(r)$ 只是 r 的函数，所以式（2-2）用球坐标表示显然更方便，直角坐标与球坐标的空间关系如图 2-1 所示，两种坐标的关系式为

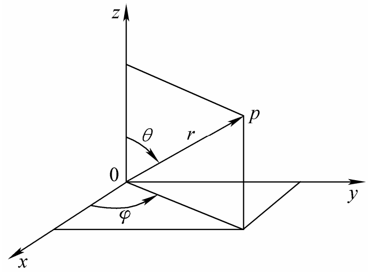


图 2-1　直角坐标与球坐标的空间关系

$$\left.\begin{aligned} x &= r\sin\theta\cos\varphi \\ y &= r\sin\theta\sin\varphi \\ z &= r\cos\theta \\ r &= \sqrt{x^2+y^2+z^2} \end{aligned}\right\} \tag{2-3}$$

在球坐标中，r 的取值范围为 0～+∞；θ 的取值范围为 0～π；φ 的取值范围为 0～2π。直角坐标系中描述的波函数 $\psi(x, y, z)$可转变成用球坐标来描述 $\psi(r, \theta, \varphi)$，式（2-2）化为球坐标方程为

$$\frac{\partial}{\partial r}\left(r^2\frac{\partial\psi}{\partial r}\right) + \frac{2m_e r^2}{\hbar^2}\left(\varepsilon + \frac{{e_s}^2}{r}\right)\psi + \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\psi}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2\psi}{\partial\varphi^2} = 0 \tag{2-4}$$

令 $\psi(r, \theta, \varphi) = R(r)Y(\theta, \varphi)$，将其代入式（2-4）中，再除以 $R(r)Y(\theta, \varphi)$，便得到

$$\frac{1}{R}\frac{\mathrm{d}}{\mathrm{d}r}\left(r^2\frac{\mathrm{d}R}{\mathrm{d}r}\right) + \frac{2m_e r^2}{\hbar^2}\left(\varepsilon + \frac{{e_s}^2}{r}\right) = -\frac{1}{Y}\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial Y}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2 Y}{\partial\varphi^2}\right] \tag{2-5}$$

式（2-5）的左边只是 r 的函数，而右边只是 θ 和 φ 的函数。在这种情况下，要使式（2-5）成立，它的两边必须等于同一个常数。设此常数为 λ，则有

$$\frac{1}{R}\frac{\mathrm{d}}{\mathrm{d}r}\left(r^2\frac{\mathrm{d}R}{\mathrm{d}r}\right) + \frac{2m_e r^2}{\hbar^2}\left(\varepsilon + \frac{{e_s}^2}{r}\right) = \lambda \tag{2-6}$$

$$-\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial Y}{\partial\theta}\right)+\frac{1}{\sin^2\theta}\frac{\partial^2 Y}{\partial\varphi^2}\right]=\lambda Y \tag{2-7}$$

解式（2-6）和式（2-7）就可以求出氢原子波函数的角向部分 $Y(\theta, \varphi)$和径向部分 $R(r)$。下面就分别求 $Y(\theta,\varphi)$和 $R(r)$。

2. 氢原子波函数的角向部分

欲求 $Y(\theta,\varphi)$，需要解方程式（2-7）。以 $\sin^2\theta$ 乘式（2-7）的两边，可得

$$\sin\theta\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial Y}{\partial\theta}\right)+\lambda Y\sin^2\theta=-\frac{\partial^2 Y}{\partial\varphi^2} \tag{2-8}$$

用分离变量的方法可以再将 θ 和 φ 分离。令 $Y(\theta,\varphi)=\Theta(\theta)\Phi(\varphi)$，并代入式（2-8）中，再除以 $\Theta(\theta)\Phi(\varphi)$，就得到

$$\frac{\sin\theta}{\Theta}\frac{\mathrm{d}}{\mathrm{d}\theta}\left(\sin\theta\frac{\mathrm{d}\Theta}{\mathrm{d}\theta}\right)+\lambda\sin^2\theta=-\frac{1}{\Phi}\frac{\mathrm{d}^2\Phi}{\mathrm{d}\varphi^2} \tag{2-9}$$

要使式（2-9）成立，它的左右两边也必须同时等于同一个常数。设这一常数为 m^2，则有

$$\frac{\sin\theta}{\Theta}\frac{\mathrm{d}}{\mathrm{d}\theta}\left(\sin\theta\frac{\mathrm{d}\Theta}{\mathrm{d}\theta}\right)+\lambda\sin^2\theta=m^2 \tag{2-10}$$

及

$$-\frac{1}{\Phi}\frac{\mathrm{d}^2\Phi}{\mathrm{d}\varphi^2}=m^2 \tag{2-11}$$

之所以将这一常数设为 m^2，是因为波函数的标准条件要求它必须为正值，如果取负值 $-k^2$，出现式（2-11）的解是 $\Phi=A\mathrm{e}^{k\varphi}$。这样，当 $\varphi=0$，±2π，±4π，…时，Φ 的值均不相等。而 $\varphi=0$，±2π，±4π，…，指的是同一角度。这等于说，对于同一角度 φ，Φ 有不同的值，也就是说 Φ 不是 φ 的单值函数。这就违背了波函数的标准条件——单值性。因而对于式（2-9）所设的常数必须是正值。于是式（2-11）的解为 $\Phi=A\mathrm{e}^{\mathrm{i}m\varphi}$，归一化后得到

$$\Phi=\frac{1}{\sqrt{2\pi}}\mathrm{e}^{\mathrm{i}m\varphi} \tag{2-12}$$

为了使 Φ 是变量 φ 的单值函数，m 必须是零或整数，即 $m=0$，±1，±2，…。

以上求出了波函数角向部分中的 $\Phi(\varphi)$。下面再来求 $\Theta(\theta)$。欲求 $\Theta(\theta)$，需要解方程式（2-10），即

$$\frac{1}{\sin\theta}\frac{\mathrm{d}}{\mathrm{d}\theta}\left(\sin\theta\frac{\mathrm{d}\Theta}{\mathrm{d}\theta}\right)+\left(\lambda-\frac{m^2}{\sin^2\theta}\right)\Theta=0 \tag{2-13}$$

为了方便起见，可作变量代换，令 $x=\cos\theta$，那么式（2-13）变为

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[(1-x^2)\frac{\mathrm{d}\Theta}{\mathrm{d}x}\right]+\left(\lambda-\frac{m^2}{1-x^2}\right)\Theta=0,\ (-1\leqslant x\leqslant 1) \tag{2-14}$$

由式（2-14）的波函数的标准条件（单值性）要求，可得

$$\lambda = l(l+1)\ \ (l=0,\ 1,\ 2,\ \cdots) \tag{2-15}$$

满足式（2-15），方程式（2-14）的解为

$$\Theta(\theta) = P_l^{|m|}(x) = (1-x^2)^{\frac{|m|}{2}} \frac{\mathrm{d}^{|m|}}{\mathrm{d}x^{|m|}} P_l(x) \tag{2-16}$$

式中，$P_l(x)$是 l 阶勒让德多项式；$P_l^{|m|}(x)$ 称为缔合勒让德多项式。它只有在$|m| \leqslant l$时才不等于零，即$|m|$不能超过 l，最大等于 l。因而，对于一定的 l 值，m 只能取下列$(2l+1)$个值。

$$m = 0,\ \pm 1,\ \pm 2,\ \cdots,\ \pm l \tag{2-17}$$

将式（2-12）和式（2-16）代入 $Y(\theta, \varphi) = \Theta(\theta)\Phi(\varphi)$中，便得到

$$Y_{lm}(\theta,\ \varphi) = N_{lm} P_l^{|m|}(\cos\theta)\mathrm{e}^{\mathrm{i}m\varphi} \tag{2-18}$$

利用归一化条件

$$\iint Y_{lm}(\theta,\ \varphi) Y_{lm}(\theta,\ \varphi)\sin\theta \mathrm{d}\theta \mathrm{d}\varphi = 1 \tag{2-19}$$

可求得

$$N_{lm} = \sqrt{\frac{(l-|m|)!(2l+1)}{(l+|m|)!4\pi}} \tag{2-20}$$

可见，只要给出 l 及 m 的值，即可求出相应的 $Y_{lm}(\theta,\ \varphi)$。表 2-1 给出 l = 0，1，2 这三种情况下氢原子波函数角向部分的具体形式。

表 2-1　波函数角向部分的具体形式

l	m	$Y_{lm}(\theta, \varphi)$
0	0	$Y_{00} = 1/\sqrt{4\pi}$
1	1	$Y_{11} = -\sqrt{3/(8\pi)}\sin\theta \mathrm{e}^{\mathrm{i}\varphi}$
	0	$Y_{10} = \sqrt{3/(4\pi)}\cos\theta$
	−1	$Y_{1,-1} = \sqrt{3/(8\pi)}\sin\theta \mathrm{e}^{-\mathrm{i}\varphi}$
2	2	$Y_{22} = \sqrt{15/(32\pi)}\sin^2\theta \mathrm{e}^{\mathrm{i}2\varphi}$
	1	$Y_{21} = \sqrt{15/(8\pi)}\sin\theta\cos\theta \mathrm{e}^{\mathrm{i}\varphi}$
	0	$Y_{20} = \sqrt{5/(16\pi)}(3\cos^2\theta - 1)$
	−1	$Y_{2,-1} = \sqrt{15/(8\pi)}\sin\theta\cos\theta \mathrm{e}^{-\mathrm{i}\varphi}$
	−2	$Y_{2,-2} = \sqrt{15/(32\pi)}\sin^2\theta \mathrm{e}^{-\mathrm{i}2\varphi}$

3. 氢原子波函数的径向部分

氢原子波函数的角向部分 $Y_{lm}(\theta, \varphi)$只能告诉我们电子出现在某一方位的几率，

也就是电子出现在 $\theta\to\theta+\mathrm{d}\theta$ 和 $\varphi\to\varphi+\mathrm{d}\varphi$ 的立体角内的几率的大小，要想知道电子出现在距原子核不同距离处的几率的大小，必须求氢原子波函数的径向部分 $R(r)$，这就需要解方程式（2-6）。

方程式（2-6）可以改写成

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{\mathrm{d}R}{\mathrm{d}r}\right)+\frac{2m_e}{\hbar^2}\left(\varepsilon+\frac{e_s{}^2}{r}\right)R=\frac{1}{r^2}\lambda R \tag{2-21}$$

将式（2-15）代入式（2-20），即得

$$\frac{1}{r^2}\frac{\mathrm{d}}{\mathrm{d}r}\left(r^2\frac{\mathrm{d}R}{\mathrm{d}r}\right)+\left[\frac{2m_e\varepsilon}{\hbar^2}+\frac{2m_ee_s{}^2}{\hbar^2 r}-\frac{l(l+1)}{r^2}\right]R=0 \tag{2-22}$$

当能量为正值时（$\varepsilon>0$），无论 ε 等于任何值，式（2-22）的解都满足波函数的标准条件，即体系的能量具有连续谱，在无穷远处波函数不为零。能量为正值意味着电子不再受原子核的约束，这对应于电离状态的电子。换句话说，当电子电离后其能量是连续可变的。然而当能量为负值时，我们看到，电子受到原子核的束缚，这时只有当 ε 取下列特定值时，方程式（2-22）的解才满足波函数的标准条件。

$$\varepsilon_n=-\frac{1}{n^2}\frac{m_ee_s{}^4}{2\hbar^2}\quad(n=1,\ 2,\ 3,\ \cdots) \tag{2-23}$$

在能量满足式（2-23）条件下，方程式（2-22）的解为

$$R_{nl}(r)=N_{nl}\mathrm{e}^{-\frac{r}{na_0}}\left(\frac{2r}{na_0}\right)^l L_{n+1}^{2l+1}\left(\frac{2r}{na_0}\right) \tag{2-24}$$

其中 $a_0=\hbar^2/(m_ee_s{}^2)$恰好是氢原子的第一玻尔轨道半径，N_{nl} 是归一化因子。利用归一化条件

$$\int R_{nl}(r)R_{nl}^*(r)r^2\mathrm{d}r=1 \tag{2-25}$$

式中，$R_{nl}^*(r)$ 为 $R_{nl}(r)$ 的共轭复数。由式（2-25）可以得到

$$N_{nl}=-\left\{\left(\frac{2}{na_0}\right)^3\frac{(n-l-1)!}{2n[(n+1)!]^3}\right\}^{1/2} \tag{2-26}$$

将式（2-26）代入（2-24）可得

$$R_{nl}(r)=-\sqrt{\left(\frac{2}{na_0}\right)^3\frac{(n-l-1)!}{2n[(n+l)!]^3}}\mathrm{e}^{-\frac{r}{na_0}}\left(\frac{2r}{na_0}\right)^l L_{n+1}^{2l+1}\left(\frac{2r}{na_0}\right) \tag{2-27}$$

其中 $L_{n+l}^{2l+1}(\rho)=\dfrac{d^{2l+1}}{d\rho^{2l+1}}L_{n+1}(\rho)$ 称为缔合拉盖尔多项式。量子数 l 的最小值为零，最大值为$(n-1)$，即 $l=0$，1，2，⋯，$(n-1)$。一般称 $l=0$ 的态为 s 态，$l=1$，2，3⋯的态依次称为 p，d，f⋯态，处于这些态的电子，依次称为 s，p，d，f⋯电子。当 n 和 l 取不同值时，氢原子的径向波函数将有不同的形式。表 2-2 给出 $n=1$，2，3 三种情况下的径向波函数。

表 2-2　径向波函数的具体形式

n	l	$R_{nl}(r)$
1	0	$R_{10}=(1/a_0)^{3/2}\,2\mathrm{e}^{-r/a_0}$
2	0	$R_{20}=[1/(2a_0)]^{3/2}(2-r/a_0)\mathrm{e}^{-r/(2a_0)}$
	1	$R_{21}=[1/(2a_0)]^{3/2}\left[r/\left(a_0\sqrt{3}\right)\right]\mathrm{e}^{-r/(2a_0)}$
3	0	$R_{30}=[1/(3a_0)]^{3/2}\left[2-4r/(3a_0)+4/27(r/a_0)^2\right]\mathrm{e}^{-r/(3a_0)}$
	1	$R_{31}=(2/a_0)^{3/2}\left[2/\left(27\sqrt{3}\right)-r/\left(81a_0\sqrt{3}\right)\right](r/a_0)\mathrm{e}^{-r/(3a_0)}$
	2	$R_{32}=(2/a_0)^{3/2}\,1/\left(81\sqrt{15}\right)(r/a_0)^2\,\mathrm{e}^{-r/(3a_0)}$

归纳上述内容可以得到氢原子的定态波函数为

$$\psi_{nlm}(r,\ \theta,\ \varphi)=R_{nl}(r)Y_{lm}(\theta,\ \varphi) \qquad (2\text{-}28)$$

它是由量子数所表征的。这三个量子数的值分别为

$$\left.\begin{aligned}&n=1,\ 2,\ 3\cdots\\&l=0,\ 1,\ 2,\ \cdots,\ (n-1)\\&m=0,\ \pm1,\ \pm2,\ \cdots,\ \pm l\end{aligned}\right\} \qquad (2\text{-}29)$$

薛定谔方程解中的量子数有十分明确的物理意义。n 为主量子数，它决定着电子运动状态的能量（负值），在氢原子或类氢离子体系中，n 相同的电子具有相同的能量。n 与能层或电子层相对应，n = 0，1，2，⋯分别称电子处于第一、二、三、⋯能层，常用光谱符号 K, L, M, N, ⋯分别表示；l 为角量子数，由 l 决定电子运动角动量的大小而得名，它决定了电子在空间的角度分布与电子云的形状。在多电子体系中 l 还影响电子的能量。l 决定着同一层中能级的大小，因而常称为副量子数。l 取值为 0，1，2，⋯，$(n-1)$分别可用相应的光谱符号 s, p, d, f, ⋯等来标记；m 为磁量子数，角量子数为 l 的电子在外磁场中有不同的取向。它在磁场方向上的分量由 m 决定。

4. 氢原子核外电子的几率分布

氢原子问题实质上是氢原子的核外电子相对于核的运动问题。$\Psi_{nlm}(r,\ \theta,\ \varphi)$ =

$R_{nl}(r)Y_{lm}(\theta,\varphi)$是描写氢原子中电子运动的波函数。因此，氢原子中电子处在核外某体积元 dτ（图 2-2）中的几率为

$$\begin{aligned} w_{nlm}(r,\theta,\varphi)\mathrm{d}\tau &= \left|\psi_{nlm}(r,\theta,\varphi)\right|^2 \mathrm{d}\tau = \left|R_{nl}(r)Y_{lm}(\theta,\varphi)\right|^2 r^2 \sin\theta \mathrm{d}r\mathrm{d}\theta\mathrm{d}\varphi \\ &= \left[\left|R_{nl}(r)\right|^2 r^2\mathrm{d}r\right]\left[\left|Y_{lm}(\theta,\varphi)\right|^2 \sin\theta\mathrm{d}\theta\mathrm{d}\varphi\right] \end{aligned} \tag{2-30}$$

上式中的 $\left|R_{nl}(r)\right|^2 r^2\mathrm{d}r$ 表示在半径为 r 到 $r+dr$ 的球壳内找到电子的几率；而 $\left|Y_{lm}(\theta,\varphi)\right|^2 \sin\theta\mathrm{d}\theta\mathrm{d}\varphi$ 则表示在立体角 $\mathrm{d}\Omega = \sin\theta\mathrm{d}\theta\mathrm{d}\varphi$ 内找到电子的几率，或者说是限定在 $\theta\to\theta+\mathrm{d}\theta$ 和 $\varphi\to\varphi+\mathrm{d}\varphi$ 范围内找到电子的几率。

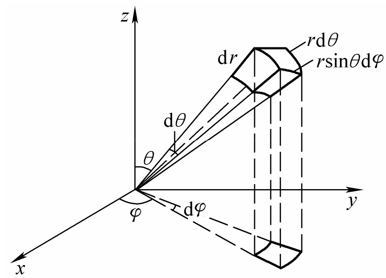


图 2-2 氢原子核外体积元

从表 2-2 可以看出，径向波函数既决定于量子数 n，也决定于量子数 l。在半径为 r 至 $r+\mathrm{d}r$ 的球壳内找到电子的几率为

$$w_{nl}(r)\mathrm{d}r = [R_{nl}(r)]^2 r^2\mathrm{d}r \tag{2-31}$$

根据这一关系式就可画出如图 2-3 所示电子径向几率分布。

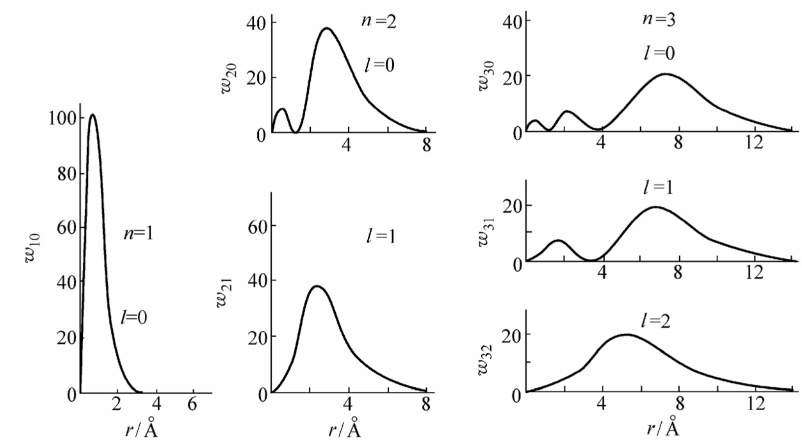


图 2-3 氢原子中电子径向几率分布（1Å =0.1nm）

对于 1s 态的电子（$n=1$，$l=0$），有

$$w_{10}(r)\mathrm{d}r = \left|R_{10}(r)\right|^2 r^2\mathrm{d}r = \left[\left(\frac{1}{a_0}\right)^{3/2} 2\mathrm{e}^{-\frac{r}{a_0}}\right]^2 r^2\mathrm{d}r = \frac{4}{a_0^{\ 3}} 2\mathrm{e}^{-\frac{r}{a_0}} r^2\mathrm{d}r \tag{2-32}$$

对于 2s 态的电子（$n=2$，$l=0$），有

$$w_{20}(r)\mathrm{d}r=\left|R_{20}(r)\right|^2 r^2\mathrm{d}r=\left[\left(\frac{1}{2a_0}\right)^{3/2}\left(2-\frac{r}{a_0}\right)\mathrm{e}^{-\frac{r}{a_0}}\right]^2 r^2\mathrm{d}r$$
$$=\left(\frac{r^2}{2a_0^{\ 3}}-\frac{r^3}{2a_0^{\ 4}}+\frac{r^4}{8a_0^{\ 5}}\right)2\mathrm{e}^{-\frac{2r}{a_0}}\mathrm{d}r \qquad (2\text{-}33)$$

可以证明 $w_{10}(r)$的最大值在 $r = a_0$ 处（a_0 为玻尔第一轨道半径），且在 $r = 0$ 及 $r = \infty$ 处 $w_{10}(r) = 0$。这说明，在 $r = 0$ 和 $r = \infty$除外的其余空间中找到电子的几率均不为零。（注意，这是指径向分布，尚未考虑角向分布）。在 $r = a_0$ 附近找到电子的几率最大。同样可以得到 2p 态和 3d 态电子的径向几率密度的最大值分别是在 $r = 4a_0$ 和 $r = 9a_0$ 处。

至于氢原子中的电子按角度的分布，由于电子位于立体角 dΩ 内的几率为

$$w_{lm}(\theta,\varphi)\mathrm{d}\Omega=\left|Y_{lm}(\theta,\varphi)\right|^2\mathrm{d}\Omega=\left|Y_{lm}(\theta,\varphi)\right|^2\sin\theta\mathrm{d}\theta\mathrm{d}\varphi \qquad (2\text{-}34)$$

因此表 2-1 已给出几个态的 $Y_{lm}(\theta, \varphi)$的具体形式，利用该表即可画出 s 态（$l = 0$），p 态（$l = 1$）和 d 态（$l = 2$）电子的几率按角度的分布，如图 2-4 所示。图中给出的是几率与 θ 角的关系。对于与竖直轴的夹角为 θ 的某一射线来说，阴影所截取射线的长度即表示与该 θ 角相对应的几率。至于 $Y_{lm}(\theta, \varphi)$与 φ 的关系，由于 $Y_{lm}(\theta, \varphi)$

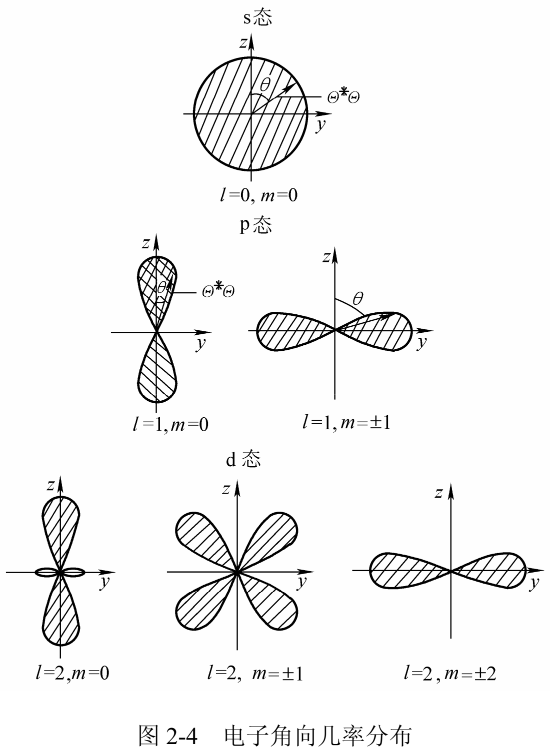


图 2-4　电子角向几率分布

中与 φ 有关的因子是 $e^{im\varphi}$，故 $|Y_{lm}(\theta,\varphi)|^2$ 与 φ 无关。于是几率按角度的分布图就是将图 2-4 中的每一个小图形分别绕竖直轴旋转半周所得到的立体图形。

原子轨函 Ψ 的径向和角度分布图只表明 Ψ 数值在空间的变化，不表明电子运动的轨迹。若要形象地理解原子轨函，应该用电子云图——$|\Psi|^2$ 在空间的分布，即电子在空间出现的几率密度的分布图。综合 $|R|^2\!-\!r$ 图与 $|Y|^2-\theta,\varphi$ 图，常用小黑点的密疏来表示 $|\Psi|^2$ 在空间的大小，如图 2-5a 所示。有时用等密度线的方式，以最密处定为 1，其他各面注以相对密度值，如图 2-5b 所示。也可用界面由来表示，此时所画出的界面表明在此界面内电子出现的几率达到 90%，如图 2-5c 所示。

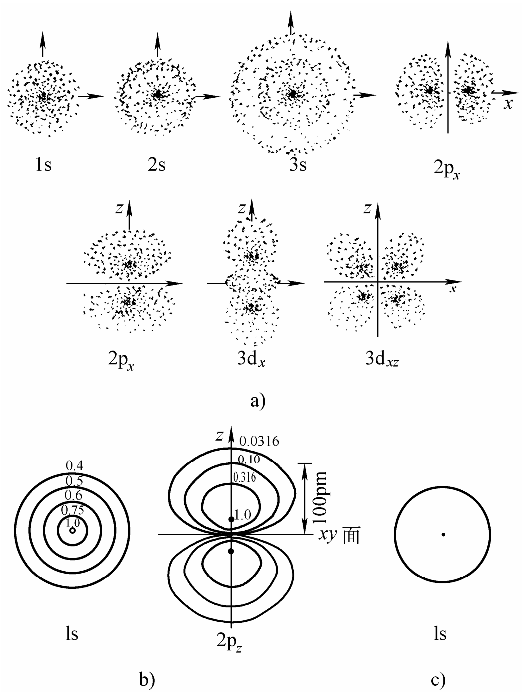


图 2-5 电子云图像的几种方式

a）电子密疏图 b）等密度图 c）界面图

5. 氢原子中电子能量

前面已导出氢原子中电子的能量为式（2-23），式（2-23）说明氢原子中电子的能量只能取一系列分立值。能级从基态$-m_e e_s^4/(2\hbar^2)$起一直延伸到零，而且愈来愈密。这是由于库仑势的绝对值随 r 增大而缓慢减小的缘故。

n 值确定之后，电子的能量也就确定了，然而，对于一个确定的 n 值，电子的波函数还可以有不同的形式。这是由于波函数 $\Psi_{nlm}(r,\theta,\varphi)$与 n，l，m 有关。这一点在表 2-1 及表 2-2 中表示得很清楚。从表中可以看出：对于一个给定的 n 值，l

可以取 0，1，2，⋯，$(n-1)$；而对于一个给定的 l 值，m 又可以取$(2l+1)$个不同的值。可见，对应于一个能级可以有 $\sum_{l=0}^{n-1}(2l+1)=n^2$ 个状态。因此，我们说氢原子中电子的能级是简并的。第 n 个能级的简并度等于 n^2。例如，当 n=3 时，氢原子中电子有一个能级，而波函数却有 9 种，它们是：Ψ_{300}、Ψ_{311}、Ψ_{310}、$\Psi_{31\text{-}1}$、Ψ_{322}、Ψ_{321}、Ψ_{320}、$\Psi_{32\text{-}1}$、$\Psi_{32\text{-}2}$。

本节所讨论的关于氢原子的性质，可以很方便地推广到类氢离子以及在有心力场中运动的粒子。对于类氢离子来说，它的原子核的电量是 Ze，这使得它的势能不再是 $U(r)=-e^2/(4\pi\varepsilon_0 r)$，而是 $U(r)=-Ze^2/(4\pi\varepsilon_0 r)$。这一差别与波函数的角向部分无关，它只对径向波函数有影响。而且只需将氢原子的径向波函数中的 $1/a_0$ 换成 Z/a_0 即可。

2.1.2　核外电子的排布规律及周期表中各元素原子的电子层结构

1. 多电子原子体系中电子运动的描述及近似能级图

除了氢原子或类氢离子之外，所有的原子（离子）都属于多电子原子（离子）。这样的体系原则上只需在薛定谔方程中补进其他电子对核的作用能及增加所有电子与电子之间的相互作用势能项，求解即可。但难以获得精确解，只能近似解出各轨道及其能量。其解具有类似于上述氢原子解的形式，不过各轨道的能量和能级次序将不同于氢原子的轨道，并且多个电子将处于可能的不同轨道上。不同轨道电子云密集于核外的距离不同，1s 电子离核最近，因此，它将有效地屏蔽原子核对其他轨道电子的作用。反之，处于其他轨道的电子对 1s 电子的屏蔽则显得不大有效，甚至无效。核外的其他电子的存在笼罩着原子核减小了核对指定电子的吸引力，这种现象被称为屏蔽效应。粗略地讲，对于多电子体系中指定的某个 j 电子而言，由于其他轨道上的电子对核的屏蔽作用，使核对 j 电子的吸引力减小，就好像核电荷 Z 减小一样，即从原来的 Z 减为 Z_j*，Z_j*称为有效核电荷，为

$$Z_j^{\ *}=Z-\sum_{i\neq j}^{n-1}\sigma_i \tag{2-35}$$

式中 σ_i 是 j 电子以外的 i 电子对核屏蔽（影响到 j 电子）的常数，通常可由某些经验规律计算出来。j 电子的轨道能量 $\varepsilon_{nl,j}$（单位为 J）为

$$\varepsilon_{nl,j}=-2.18\times10^{-18}(Z_j^{\ *}/n)^2 \tag{2-36}$$

此外，多电子原子中 n 较大的原子轨道，如 4s 轨道的电子有相当的几率出现在核附近，好似钻入内部，部分地回避了（$n-1$）层的原子轨道，如 3d 轨道的电子对它的屏蔽作用，却又反过来削弱了 3d 电子受核的吸引力。这就导致了多电子原子体系中 n 较小的 3d 电子的能量略高于 n 较大的 4s 电子的能量。这种现象称

为钻穿效应，这样又进一步说明了多电子原子各电子的能量应由主量子数 n 与角量子数 l 共同来决定。

我国化学家徐光宪院士提出了描述多电子体系的原子轨道近似能级次序的（n+0.7l）规则：以该轨道的（n+0.7l）数值大小决定轨道能量之高低，并将（n+0.7l）值的第一位数字相同的各能级编成一组称为能级组，各能级组内各个轨道能级差较小，而能级组间能量差较大，外层电子所在能级组的编号恰好是化学元素所在的周期数。这样，多电子原子轨道的能量由低到高依次为：1s；2s, 2p；3s, 3p；4s, 3d, 4p；5s, 4d, 5p；6s, 4f, 5d, 6p；7s, 5f, 6d, 7p；…参见图 2-6。多电子原子轨道的能量顺序反映了：l 相同时，$\varepsilon_K < \varepsilon_L < \varepsilon_M < \varepsilon_N < \cdots$；$n$ 相同时，$\varepsilon_{ns} < \varepsilon_{np} < \varepsilon_{nd} < \varepsilon_{nf}$；$n$ 和 l 不同时，发生能级交错，但遵循 $\varepsilon_{ns} < \varepsilon_{(n-2)f} < \varepsilon_{(n-1)d} < \varepsilon_{np}$。此外，不同元素原子相应原子轨道的能量会有很大差别。

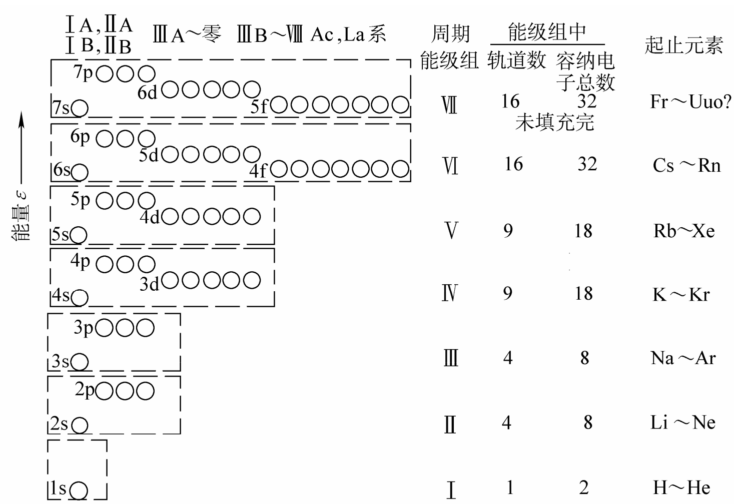


图 2-6　原子轨道近似能级、能级组与周期律

2. 基态原子核外电子排布遵循规律

原子处于基态时，其核外电子的排布应遵循三条重要原则。

1）泡利不相容原理　一个原子轨道最多只能容纳两个自旋相反的电子，或者在一个原子中不可能存在 4 个量子数（除了上面提到的三个量子数，还有自旋量子数 m_s，为±1/2）完全相同的两个电子。

2）能量最低原理　在不违背泡利原理的条件下，电子优先占据能量较低的原子轨道，即电子按低能轨道顺序向高能轨道逐一填充，使整个原子体系能量处于最低状态。

3）洪特规则　在等价轨道上的电子尽可能以相同的自旋方向分占不同的轨

道；当等价轨道上电子处于全充满（p^6 或 d^{10} 或 f^{14}）、半充满（p^3 或 d^5 或 f^7）时，原子能量较低，比较稳定。

根据上述三原则，可以写出各元素基态原子的电子结构，例如 Fe，原子序 $Z=26$，则核外 26 个电子的构型或排布可表示为 $1s^2\,2s^2\,2p^6\,3s^2\,3p^6\,4s^2\,3d^6$。化学反应时，参与反应的只是原子的外层电子。其内层电子结构通常是不变的。不参与反应的内层电子称为原子实。可以用相应的稀有气体加方括号来表示原子实。如[He]表示了 $1s^2$ 的两个电子在最内层电子轨道上的排布，而[Kr]表示了 $1s^2\,2s^2\,2p^6\,3s^2\,3p^6\,4s^2\,3d^{10}\,4p^6$ 36 个内层电子的结构。参与化学反应的外层电子是填充在最高能级组中的有关轨道上的电子，这组轨道常常称为价电子轨道或称价轨道。占有这些轨道的电子通常就称为价电子。于是任何元素原子皆可以用：[稀有气体]价电子的方式来表示。Fe 的电子结构应该表述为 Fe[Ar] $4s^2\ 3d^6$。对于基态 Cr 原子应表示为 Cr [Ar] $4s^1\ 3d^5$，而基态 Cu 原子则为 Cu [Ar] $4s^1\ 3d^{10}$。根据洪特规则，半充满 d^5 及全充满的 d^{10} 电子排布比较稳定，因此 Cr 与 Cu 在 4s 轨道上只占有 1 个电子，这为光谱实验数据所确证。至于 Sn（$Z=50$）的电子构型应表示为 Sn [Kr] $5s^2\,4d^{10}\,5p^2$。这里的 $4d^{10}$ 并不包括在 Kr 的电子构型之中。

当原子电离成为离子时，丢电子的顺序应满足 $n+0.4l$ 规则，即此值越大的电子应先电离。因此在 ns $(n-1)$d np 组成的能级组中，先电离的是 np 电子，然后是 ns 电子，而$(n-1)$d 电子最后失去。Cu 原子电离时先丢失的是 4s 电子，然后丢失 3d 电子。这样 Cu^{2+}的电子结构应表示为 Cu^{2+} [Ar] $3d^9$。

3. 元素周期表与原子的电子层结构

根据元素化学和物理性质的相似性分组并按性能周期性变化规律进行排列形成的元素周期表无疑是化学史上的一个重要的里程碑。随着对原子结构研究的不断深入，逐步揭示了原子核外电子排布的规律，从而指明了元素周期律的内在原因。元素周期律是原子中电子层结构周期性变化的必然结果。分析一下周期表（参见表 2-3）可得出如下关系。

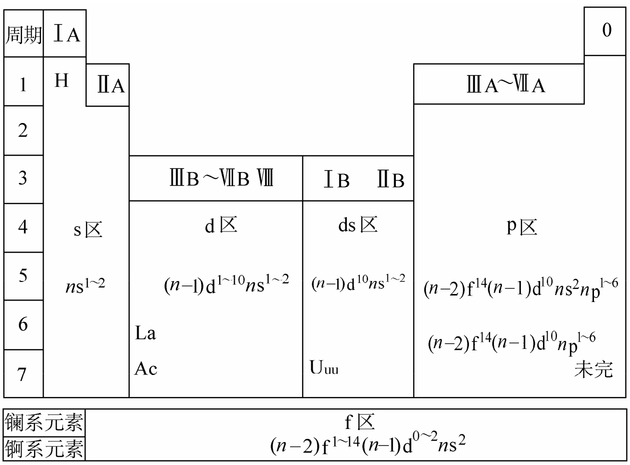
① 元素所属周期是由其基态原子能量最高的电子所在的能级组的序号所规定。各周期中元素数目是相应能级组中原子轨道可容纳的电子总数。每一周期都以 $n\mathrm{s}^1$ 开始，以 $1s^2$（He）或 $n\mathrm{p}^6$（Ne、Ar、Kr、Xe、Rn）为终结。因此周期表中含有一个特短周期（2 种元素），两个短周期（各 8 种元素），两个长周期（各 18 种元素），一个特长周期（32 种元素）和一个未完成周期（理论推测应有 32 种元素，1994 年 12 月已人工合成了第 111 号元素）。长周期中各有 10 种副族元素（记以 B）——过渡元素，特长周期中各含有 15 种内过渡元素（镧系与锕系元素）。

② 主族（ⅠA～ⅦA）及副族（ⅠB～ⅡB）的族数是基态原子最高占有能级组中（ns+ np）电子数，但电子构型为 $1s^2$ 的 He 及 $n\mathrm{s}^2 n\mathrm{p}^6$ 的稀有气体（Ne，Ar，…）

称零族。副族ⅢB～ⅦB 族数是基态原子最高占有能级组中[ns+ (n–1)d]的电子数。最高占有能级组中 $ns^2(n-1)d^{6\sim8}$ 的电子构型有 3 纵列，称Ⅷ族。共有ⅠA～ⅦA、ⅠB～ⅡB、Ⅷ和零 16 个族。

③ 根据原子的电子层结构特征，可把全部元素分成 5 个区，即 s，p，d，ds 和 f 区。s 区包括ⅠA 和ⅡA（碱金属，碱土金属），电子层构型是 $ns^{1\sim2}$；p 区包括ⅢA～ⅦA 及零族元素，电子层构型是 $np^{1\sim6}$；d 区包括ⅢB～ⅦB 及Ⅷ族，电子层构型为 $ns^{1\sim2}(n-1)d^{1\sim8}$；ds 区包括ⅠB 和ⅡB 族，电子层构型是 $ns^{1\sim2}(n-1)d^{10}$；f 区为 La 系、Ac 系，各 14 种元素，电子构型为 $(n-2)f^{1\sim14}\,(n-1)d^{0\sim2}\,ns^2$。因此，已知元素原子序数即可推知其在周期表中的位置。

表 2-3　原子价电子构型与周期系分区



4. 元素周期性

原子结构决定着元素的性质，决定各元素原子结构特征的有三个重要因素：核外电子构型，原子核有效核电荷和原子半径。随着原子序数的递增，元素性质逐渐发生质变。但各元素性质间既存在着本质差异，又有内在联系并呈现周期性，因此元素的原子的半径、电离能、电负性和金属性等化学、物理性质有着一定的渐变规律。正因为如此，根据周期表中的位置往往可推知关于元素的许多重要信息。

（1）核外电子构型

元素的化学性质很大程度上取决于价电子构型。凡基态呈稀有气体构型，即填满 p 电子（np^6）构型者是稳定结构。这种结构的原子具有相当的化学惰性。其他结构的原子或获取或丢失电子达到稳定结构而呈现化学活性。各族元素外层能级都有相似的电子构型。因而价态及化学性质相似。其中过渡元素，电子除填入 $(n-1)$d 或（和）$(n-2)$f 轨道上外，常还有一两个更易丢失的 ns 电子，因而它们既

有共同的价态，又有各自的多种氧化态，可形成多种价态的化合物，而且它们常各自呈现独特又美丽的颜色。

（2）原子半径

关于原子半径 r，按近代原子结构的概念，核外电子呈几率分布，因而原子的大小无明确界限。但在单质和化合物中，元素的原子常以化学键结合在一起，因此原子的半径实际上根据测定它们的核间距离的实验结果推导而得。通常将同种相邻原子形成单键键距之半定为共价半径，在金属晶体中则称金属半径。

周期表中同周期内的主族元素自左向右核电荷逐渐增加，各元素的最后一个电子都填充在最外层上，由于同层上电子屏蔽较弱，因此有效核电荷是明显增加的，从而导致了原子半径的明显变小；若自上而下按族递增时，增加了电子层，而电子构型基本不变，因内层电子对外层电子的屏蔽较有效，因此有效核电荷的增加不显著，原子半径显著增大。副族元素自左向右递变时，各元素的最后一个电子都填充在$(n-1)$层上，由于较内层电子对外层电子屏蔽较强，因此有效核电荷增加不明显，因而大体上原子半径变化不大。但当次外层的 d 轨道全充满时由于$(n-1)d^{10}$的较大的屏蔽作用而导致原子半径突然明显增大。副族自上而下，有效核电荷增加不明显，因而原子半径基本不变。至于镧系元素从左至右，各元素的最后一个都填充在$(n-2)$层上，由于内层电子对外层电子的屏蔽较有效，因此有效核电荷增加很少，原子半径略有收缩（约 1pm），但累计有 14 个元素导致镧系原子半径相近，加上电子构型相似，因而镧系 15 个元素（常还包括钇）化学性质相近。镧系元素原子半径的收缩（镧系收缩），又导致其后的元素与其相应上一周期的同族元素的原子半径非常接近，如 Zr 与 Hf，Nb 与 Ta，Mo 与 W 等，它们的化学性质也极相近，常以共生矿在地球共存，化学上分离它们具有一定难度。

（3）电离能与电负性

在化学反应中，各元素原子将得到或失去电子使自己的外层电子构型变成稳定的构型，其得失电子的能力可用电离能（I）与电负性（X）来描述。

电离能（I）指基态气态原子或离子失去一个电子形成其相应的气态离子所需最低能量。同一周期中，自左向右元素的失去第一个电子的第一电离能 I_1 总体上由小变大，到稀有气体时达到最大值，但会出现有规律的曲折变化，这与电子处于充满或半充满时（s^2，p^3，p^6，d^5，…）的构型较为稳定有关，如 $I_1(B) < I_1(Be)$，$I_1(O) < I_1(N)$等。同一主族内自上而下元素的 I_1 递减。

一个原子既有得电子能力又有失电子能力，当它处于指定分子中时究竟倾向于得电子还是失电子呢？为统一说明原子在分子中的行为，鲍林（L. Pauling）综合了原子得失电子的能力，提出了元素电负性的概念。电负性是元素的原子在分子中吸引成键电子的能力。鲍林比较了 A 与 B 两原子间的生成热和 A—A，B—B 的键能数据，且指定 F 的电负性 X =4.0，从而依此得出了其他元素的电负性。电

负性越大的元素的原子在分子中吸引成键电子能力越大。在周期表中总变化趋势是：同一周期元素，从左到右电负性加大，同一族元素，从上到下电负性递减。因此，周期表中的左下角元素电负性最小，右上角则最大，其余元素居中。电负性 X 的周期性变化如图 2-7 所示。

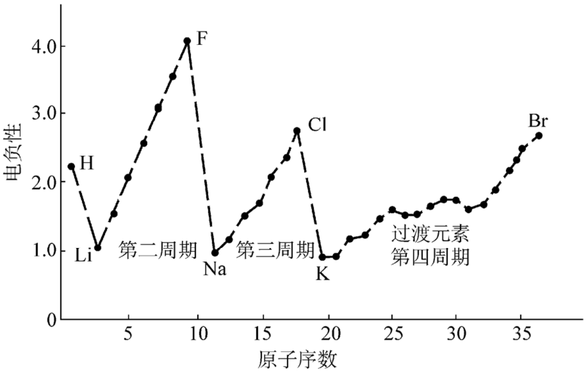


图 2-7　元素电负性与原子序数的关系

（4）金属性

元素的金属性和非金属性是指其原子在化学反应中失去和得到电子的能力。自然，比较电负性 X 的数值可判断元素的这个属性。凡 X 越小，元素的金属性越强；X 越大，非金属性越强。因此，周期表左下角与右上角元素分别是最活泼的金属与非金属。其分界限在 B，Si，As，Te，At 与 A1，Ge，Sb，Po 两条对角线元素上。此区域及其附近元素常成为半导体材料，这些元素有时称为半金属，它们的电负性约在 2.0 左右，在不同条件下或呈金属性或呈非金属性。

2.2　分子结构与化学键

对原子核外的电子构型及其运动规律有所了解后，自然要追究原子究竟如何结合成各种各样的化合物分子或晶体的问题。这就涉及化学键的问题。总起来说，原子是依靠化学键将原子或离子间的距离约束在几十至几百皮米（pm）的范围内，从而形成了分子和晶体。

由于各种原子的结构不同，不同原子相互作用时，静电的吸引和排斥的情况就不一样，不同分子中各原子相互结合的方式就不同。因此，化学键有各种不同的类型，有离子键、共价键、金属键和分子键。尽管各种键的成因有本质差别，但由于离子间的极化作用及分子的极性，实际上化学键之间没有绝对的、严格的界限，化合物分子内存在着离子型与共价型的过渡型键。

2.2.1　原子间作用力及由来

1. 原子间的作用力

两个原子间的化学亲和力，从物理本质上说，是由于在这两个原子之间存在吸引力。然而即使在化学亲和力很强的情况下，原子也难以过于接近，因为当原子相距很近时，它们之间主要显示排斥力。这种排斥力是一种短程力，即在远距离处它的数值很小，但在近距离处它比吸引力大。在平衡状态时，原子间的吸引力和排斥力相等。这时两个原子间的距离即为平衡距离。物体所受的合力为势能随距离的改变率。因此，原子在平衡距离处的势能具有极小值。

一个材料内部的势能等于每个原子的能量及相互作用能量的和。势能的零点可以任选，令原子相距很远，都处于基态，相互没有作用时，其势能 $U=0$。那么要形成一个固体，必须把这些原子从远处移到一起，假如所形成的原子集团，要比原来分散的原子更稳定，则其势能必须小于零，就是说总的相互作用能必须是负值。由于原子在吸引时做功，所以吸引的势能 U_a 是负值，它是原子间距离 r 的幂函数。$U_a=-\alpha/r^n$，α 为正的比例常数。当距离很远时，$U_a=0$，以后随距离缩短而 U_a 绝对值减小，如图 2-8a 最下面的虚线所示。相反，排斥的势能 U_r 是正值，要把互相排斥的原子聚在一起，外界必须对它做功。U_r 的大小也是原子间距离 r 的幂函数幂。$U_r=\beta/r^m$，如图 2-8a 上面虚线所示，β 也为正的比例常数。当 r 减小到一定程度时，U_r 急剧增加，致使 U_r 比 U_a 变化得更快。总的势能 U 是该两项值的和，即

$$U=U_a+U_r=-\frac{\alpha}{r^n}+\frac{\beta}{r^m}\tag{2-37}$$

U 在图 2-8a 中以实线表示。

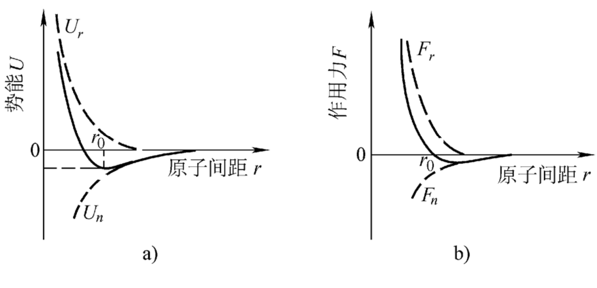


图 2-8　势能及作用力与原子间距离关系

a）势能曲线　b）作用力曲线

原子间的作用力 $F(r)$ 可由式（2-37）微分导出

$$F(r)=-\frac{dU}{dr}=-\frac{n\alpha}{r^{n+1}}+\frac{m\beta}{r^{m+1}}\tag{2-38}$$

式中，$m > n$，负项为引力 F_a；正项为斥力 F_r，见图 2-8b。当达到平衡间距 r_0 时，吸引力和排斥力平衡，即 $F(r_0) = 0$，则

$$r_0 = \left(\frac{n\alpha}{m\beta}\right)^{1/(n-m)} \tag{2-39}$$

$r = r_0$ 时，原子集团的势能具有最低值，即在图 2-8a 中如果增加或减小原子间距离，都使势能增加，原子集团变得不稳定。对于某些离子晶体，$n = 1$，$m = 9$；对于某些金属，$n = 1$，$m = 3$；对于其他类型的晶体，n 和 m 又有不同的数值。

可以设想固体的一个分子是这样的：在它中间有两个以上的原子核，外面绕有负电荷的电子云，靠静电引力结合起来。其结合强度取决于原子核之间各点的电子云密度。

2. 原子间作用力的由来

当原子由远处移近时，在原子之间产生引力。在原子间距达到几个 Å（1Å = 0.1nm）时，产生了平衡引力的斥力。斥力的产生最主要是由于原子核外面的电子云有了重叠。当两个电子壳层开始重叠时，每一壳层中的电子都不能再认为是仅属于其原来的原子了，而为分子整体所共有。在这些壳层中的电子的运动和能量仍然遵守量子规律，特别是不相容原理。如果这两个电子壳层原来都是填满的，那么当它们结合成一个统一的电子集团时，不能保证所有的电子都占据低能量的量子状态，其中有些电子就要攀登到较高的能级。这种能量的增加就引起了斥力。满壳层重叠时必然导致上述的斥力。在价电子的电子云没有重叠以前，中性原子间不会有很强的化学结合力。

当两个原子相互移近时，距原子核最远的价电子比其他电子更早发生重叠，因为价电子只占据部分填充的壳层，当它们重叠时就会产生引力。但是，当两个原子满壳层的电子云也发生重叠时，就会出现满壳层间的斥力。

价电子重叠所产生的引力，可以用氢分子的形成来说明。氢分子由两个质子和两个电子所组成。在分析电子的运动时，质子可以认为是静止的，其作用力主要考虑粒子之间最强的静电力。如果这些粒子之间的力要形成化学键，则两个电子必须把大部分时间用在两个核之间。当两个电子互相接近时，它们之间的斥力是大的；但同时，由于这两个电子同时接近两个质子，因而产生四个引力，这四个引力足够抵消上述斥力。用量子力学处理这个问题，可以知道电子在分子的各个地方运动所用去的时间，从而计算出这四个粒子之间静电力的平均值。如果电子把大部分时间用在分子中的某个地方，那么，在计算平均作用能量时，这个地方的电子与其他粒子作用的能量就较大。

电子在分子中的分布是遵守量子力学规律的。假定最初原子相距很远，相互间没有作用，这时每个原子有一个电子处于(1s)状态。这个状态的波函数是一个以

原子核为中心的球形分布的电子云。电子云密度随其与原子核的距离的增加而急剧减小。如把两个原子彼此移近，使它们的电子云恰好有不大的重叠，那么这时电子运动的特点是，电子能够从一个原子跑到另一个原子上去，而且不可能区分哪个电子是属于哪个原子。这样就改变了原来的波函数，原来是两个(1s)函数，现在需要用新的波函数来表示电子在分子整体中的运动。假设两个(1s)电子云开始重叠时没有什么变化，那么就有两个不同的(1s)型的 ψ 函数，以原子核 a 为中心有 ψ_a，以原子核 b 为中心有 ψ_b。两个电子用 1 和 2 表示，并且以 ψ_{a1} 表示电子 1 处于 ψ_a 的量子状态。这样，在分子中某点 P 的 ψ_{a1}^2 表示当电子 1 处于原子核 a 的 1s 状态时，它在 P 点出现的几率。同样，ψ_{b2}^2 表示当电子 2 处于原子核 b 的(1s)状态时，它在某点 Q 出现的几率。因此，当电子 2 在 Q 点时，在 P 点找到电子 1 的几率为它们两者几率的乘积 $\psi_{a1}^2\psi_{b2}^2$，这称为两者的合成几率。这就有可能把 $\psi_{a1}\psi_{b2}$ 看作是两个电子的总的波函数。也可设想电子 2 在 P 点有波函数 ψ_{a2}，电子 1 在 Q 点有波函数 ψ_{b1}，那么就得到总的波函数为 $\psi_{a2}\psi_{b1}$。显然，以上得到的两个波函数没有一个能单独满足要求，因为每个电子都能从一个原子移往另一个原子，并不能区分那个电子是属于那个原子的。我们所能知道的，只是当一个电子在 Q 点时，另一个电子在 P 点出现的几率，因此 $\psi_{a1}\psi_{b2}$ 与 $\psi_{a2}\psi_{b1}$ 必须有相同的机会出现。波函数的性质之一是可以把已知的波函数相加得出新的波函数。对于薛定谔方程来说，已知解的线性组合，仍为其解。因此得出新的波函数 ψ_+和 ψ_-为

$$\psi_+ = \psi_{a1}\psi_{b2} + \psi_{a2}\psi_{b1} \tag{2-40}$$

$$\psi_- = \psi_{a1}\psi_{b2} - \psi_{a2}\psi_{b1} \tag{2-41}$$

在 ψ_+中把标号 1 和 2 互换，不改变波函数；在 ψ_-中交换 1 和 2 的标号，则 ψ_-为负值。所以 ψ_+是一个对称的波函数，而 ψ_-是一个反对称的波函数。

相应于式（2-40）和式（2-41）波函数的合成几率为

$$|\psi_+|^2 = (\psi_{a1}\psi_{b2})^2 + (\psi_{a2}\psi_{b1})^2 + 2\psi_{a1}\psi_{a2}\psi_{b1}\psi_{b2} \tag{2-42}$$

$$|\psi_-|^2 = (\psi_{a1}\psi_{b2})^2 + (\psi_{a2}\psi_{b1})^2 - 2\psi_{a1}\psi_{a2}\psi_{b1}\psi_{b2} \tag{2-43}$$

在式（2-42）和式（2-43）中，前两项的四个 ψ 函数是描述一个电子围绕一个原子核处于(1s)状态运动的波函数，它们的数值在两个原子核的中心处大致是相等的。两个方程的最后一项代表两个原子的交换作用，这是因为它们有交换各自电子的可能性。交换作用对分子能量的贡献通常称为交换能。所以在两个原子核的中心处 $|\psi_+|^2$ 的值很大，而 $|\psi_-|^2$ 等于零。因此 ψ_+所表示的是电子在分子中以大部分时间用在原子核之间，并且把它们联结在一起，形成一个稳定分子的量子状态；而 ψ_-则表示电子避开这个中心区域，不能形成结合键的一个状态。原来的(1s)能级就分裂成如图 2-9 所示的两个能级 $\varepsilon_A(r)$及 $\varepsilon_s(r)$。其中 $\varepsilon_s(r)$这个能级属于 ψ_+，$\varepsilon_s(r)$曲线在 $r = 0.87$Å（0.087nm）处有一个能量最低点，说明两个氢原子接近到这种程

度时可以形成稳定的 H_2 分子，这就是结合状态（图中虚线为实验测定值），这时它们的两个电子是自旋相反的。当两个氢原子相互接近时，如果它们的电子自旋是平行的，那么原子间的相互作用永远是排斥的，如图 2-9 中 $\varepsilon_A(r)$曲线，不能形成稳定的分子。$\varepsilon_A(r)$这个能级属于 ψ_-。

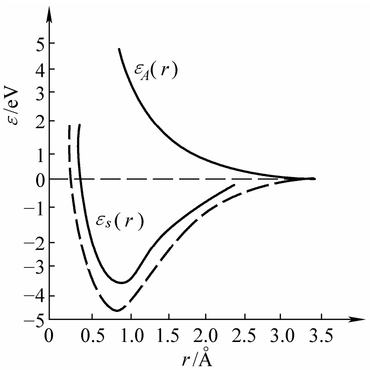


图 2-9　氢分子 1s 能级分裂

上面用氢原子价电子重叠形成氢分子解释了原子作用力的由来。这个原理同样可以用来说明金属理论中的一些问题。例如两个氢原子彼此接近时，1s 能级分裂为两个能级的情况与金属价电子能级展宽是很相似的。氢分子只有两个氢原子，1s 只能分裂成两个能级；而在金属晶体中，每立方厘米体积中具有 10^{23} 数量级个原子，因此可分裂成间隔很小数量很多的能级，称为能带，但它们的本质过程是一样的。

2.2.2　化学键

1. 离子键

在一定条件下，活泼金属原子和活泼非金属原子相互接近时，前者将失去最外层电子而形成带正电荷的阳离子；后者将获得电子而形成带负电荷的阴离子。阳离子和阴离子之间将由于静电引力而互相吸引，但当它们充分接近时，就会因离子的电子云重叠而产生斥力。当吸力和斥力相等时，可形成稳定的离子键。例如钠原子和氯原子在一定条件下互相接触时，钠原子有一个 $3s^1$ 价电子处于稳定的八个电子所组成的壳层以外，而氯原子有七个价电子 $3s^2p^5$，需要一个电子以组成八个电子的稳定壳层。因此，钠原于将失去一个电子而成为带正电荷的 Na^+离子，氯原子将夺得一个电子而成为带负电荷的 Cl^-离子。这两种离子由于相互间的静电引力而彼此吸引，直至引力被电子云重叠产生的斥力所平衡为止，形成离子化合物。离子键的特点是没有方向性和饱和性。例如，一个 Na^+离子可以同时和几个 Cl^-离子相结合，一个 Cl^-离子也可以同时和几个 Na^+离子相结合，而周围的 Cl^-离子或 Na^+离子的数目和排列方式则取决于所形成晶体的密排形式。在氯化钠晶体颗粒中，并不存在单个的氯化钠分子，化学式 NaC1 只代表氯化钠中 Na^+和 Cl^-离子的数比，并不代表分子。所以离子结合没有饱和性。另外离子的电荷分布是球形对称的，它在各个方向都可吸引电荷相反的离子，所以离子结合没有方向性。

离子型晶体的结合较强，破坏这种化学键需要较高的能量。所以离子型晶体的熔点、硬度、强度都很高，热膨胀系数很小。例如用作耐火材料的 MgO、$A1_2O_3$ 和钢中的一些非金属夹杂均属此类化合物。

离子型化合物的性质很大程度上取决于离子的性质，如离子的电荷、离子的半径、离子的电子层结构及晶体结构等。

2. 共价键

共价键理论主要有现代价键理论（valence bond theory）和分子轨道理论（molecular orbital，MO）。现代价键理论强调了分子中相邻两原子间因共享配对的电子而成键，能相当直观并成功地解释物质的化学性质，预测物质的分子结构。但由于过分强调了两原子间的电子配对，而显示出它的局限性，例如小分子 B_2、O_2 等的顺磁性价键理论无法解释。分子轨道理论自然地、合理地解释了这些现象。

现代价键理论认为：电负性差异甚大的原子间靠离子键结合成分子，电负性相近甚至相同的原子是靠原子共享电子，形成共价键而结合成分子的。H_2、H_2O、NH_3 等单质分子和 HCl、H_2O、NH_3 等化合物分子是由共价键结合的。它们是通过价电子共有，形成稳定的电子满壳层（如 2、8、18…个电子）的方式结合的。共价键的本质是原子共享电子，被共享的电子就像一个带负电荷的桥，把两个带正电荷的核吸引在一起。现代价键理论的理论要点如下。

1）配对成键　原子 A 和 B 各有一个未成对电子且自旋相反，可互相配对构成共价单键。若 A 与 B 各有两个或 3 个未成对电子，则能两两配对构成共价键或叁键，H－H，蒸气 Li－Li 分子为单键，N≡N 分子为叁键等。H(1s)－C1(3p)可形成单键，而两个 He 原子不能形成分子，因为氦原子没有未成对电子。

2）饱和性－原子价数的确定　两原子中自旋相反的电子配对后，不能再与第三个原子配对一起形成共价键。若 A 原子有两个未成对电子，而 B 原子只有一个，则 A 能与两个 B 化合成 AB_2 分子，如 H_2O。这些就是共价键饱和性的具体体现。共价键的饱和性是由“配对成键”所决定的。通常一个原子所含有的可以与其他原子配对的未成对电子数目是其原子价数。

3）方向性－电子云最大重叠原理　两原子的价电子发生重叠愈多，体系能量降低愈多，所形成的共价键愈稳固。共价键形成时尽可能采用电子云密度重叠最大方向，这是共价键具有方向性的原因。溴分子（Br_2）形成时，溴（Br）有一个不成对的 3p 态电子云，两个溴原子的两个 3p 态电子云只有在“迎头”重叠时，才发生最大有效重叠而成键，如图 2-10 所示。

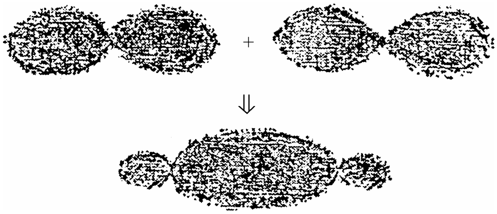
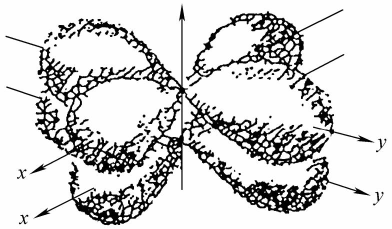


图 2-10　成键轨道的有效重叠方向

4）配位体　若原子 A 具有已配对的孤对电子，而 B 具有能量与之相近的空原子轨道，则 A 的孤对电子所占有的原子轨道与 B 的空轨道能有效地重叠，孤对电子为 A 与 B 共享，构成共价配键或配键，这样 A 被称为给电子体，而 B 则为受电子体。

5）成键方式　两个原子的原子轨道（价轨道）以“头碰头”方式重叠，形成沿键轴呈圆柱形对称的电子云，这种键称为 σ 键。例如，两个氢原子的 ls 电子形成 H_2；一个氢原子的 1s 电子与氯 3p 电子形成 HCl 分子；两个氯原子的 3p 电子的重叠形成 Cl_2。不同原子的原子轨道，如 p 轨道，以“肩并肩”方式重叠成键，这种键称为 π 键。π 键电子云有一个通过键轴几率密度为零的平面（节面），并对称分布于该节面上下，但符号相反（反对称）。如乙烯 $CH_2{=}CH_2$ 分子中两个碳的 2p 电子可构成一个 π 键。不同原子的价轨道，如 d 轨道以“面对面”的方式重叠成键，这种键称为 δ 键，见图 2-11。δ 键电子云沿两个通过键轴的节面呈对称分布。如 $Re_2Cl_8^{2-}$中，两个金属 Re 的 $d_{x^2-y^2}$ 轨道沿着 z 轴（Re－Re 键轴）方向的重叠。一般共价单键由 σ 键组成；双键由一个 σ 键与一个 π 键构成，如乙烯中的 C=C；叁重键则由一个 σ 键与两个 π 键构成，如乙炔中的 C≡C 键；四重键则由一个 σ 键、两个 π 键与一个 δ 键构成。键的多重性越大，总键能越大，键长将越短。成键 σ 电子云密集在两原子之间，键较强，而 π 和 δ 键电子云不然，它们的流动性较大，能量较高，较 σ 键易断裂，反应活性较高，因此，在多重键上易发生加成反应。



$d_{x^2-y^2}-d_{x^2-y^2}$

图 2-11　δ 键电子的共享方式

6）杂化轨道　实验证明，甲烷分子（CH_4）具有对称性很高的正四面体的空间构型，分子内 4 个化学键是均等的，有很高的化学稳定性，分子偶极矩为零。这些事实用原子的未成对电子配对形成共价键的理论是无法解释的。为此逐渐发展起来了杂化的概念和杂化轨道理论，以下介绍这一理论的要点。

原子在成键时，先受到微扰——周围环境对它的某种作用而受到激发，以致使原子的某些能量相近的轨道可以重新线性组合成能量相等的一组新轨道——杂化轨道。原子轨道的线性组合称为杂化。具体地说，同一能级组或者同一能层的轨道如 ns 与 np 轨道或 ns，np 与 nd 轨道或$(n-2)$f，$(n-1)$d，ns 与 np 轨道可按不

同方式杂化。杂化后体系总能量将略高于线性组合的平均值。但杂化后的轨道改变了原来轨道的电子云空间角度分布，因而改变了原子成键的方向性，这是分子呈现一定几何构型的原因所在。不仅如此，杂化还可以提高原子轨道的成键能力。常遇到的杂化类型有 sp 型、spd 型及 dsp 型。sp 型又分为 sp、sp^2 和 sp^3 杂化；spd 型又分为 sp^3d 和 sp^3d^2 杂化，是由 *n*s、*n*p 和 *n*d 原子轨道组成，主族元素常用其空的能量稍高 d 轨道参与杂化；dsp 型则分为 dsp^2、dsp^3、d^2sp^3 杂化，是由$(n-1)$d、*n*s 和 *n*p 轨道组成。副族元素常用未充满的$(n-1)$d 轨道参与杂化。

现在，具体看一下 sp^2 杂化。以 BF_3 为例，其中心原子硼的价轨道为 $2s^2p^1$。在形成 BF_3 过程中，B 原子的一个 2s 电子被激发到一个空 p 轨道上，然后硼原子一个 2s 轨道与两个 2p 轨道杂化，形成了 3 个 sp^2 杂化轨道。杂化过程示意如图 2-12 所示。

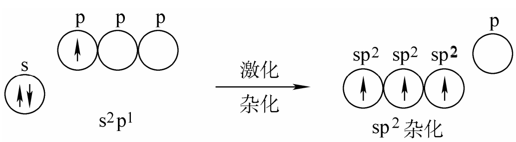


图 2-12　sp^2 杂化

这 3 个 sp^2 杂化轨道互成 120° 的夹角并分别与 F 原子的 p 轨道重叠，电子配对形成 σ 键，因而构成平面三角形分子，如图 2-13 所示。

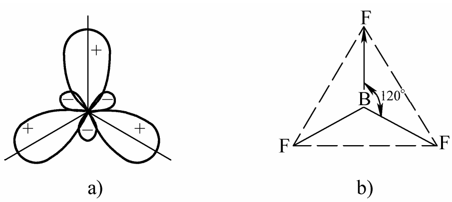


图 2-13　杂化轨道角度分布示意图和 BF_3 分子的空间构型

a）3 个 sp^2 杂化轨道角度分布　b）平面三角形结构的 BF_3 分子

分子轨道理论（MO）它可看成是量子力学线性变分法求解 H_2^+体系电子运动规律结果的推广。MO 理论与 VB 理论是从不同侧面探索研究分子结构的方法。在许多问题上，这两个方法可得出一致结论。但有时则两法各有其优缺点。因此目前两种理论尚不可偏废。不过由于计算机等新技术的兴起、迅猛发展及广泛应用，MO 理论方法已成为当代研究分子结构的最普遍和最基本的理论方法。MO 理论的基本要点如下。

1）分子中电子运动的整体性　原子中的电子在原子核为中心的力场中运动，可用原子轨道（atomic orbital，AO）波函数 ψ 来描述。分子轨道理论认为分子中的电子在整个分子的势场中运动，因此它们属于整个分子，用分子轨道波函数来

描述，记以 $\varPsi$。类似于原子轨道，分子轨道也处于一系列分立的能级上。AO 以主量子数来划分能层，用角量子数来划分能级，并用 s，p，d，f，…来标记。MO 则用希腊文 σ，π，δ，…符号来标记其在键轴方向上角动量分量（量子化的）的大小（分别为 1，2，3），并编以序号加以区分。$|\varPsi|^2 \mathrm{d}\tau$ 同样表示电子在空间小体积 $\mathrm{d}\tau$ 内出现的几率，而 $|\varPsi|^2$ 是几率密度。

2）MO 的构成 MO 由分子中各原子的原子轨道线性组合而成。其组合系数由量子力学变分法等确定。MO 的数目等于组成分子的各原子的原子轨道 AO 数目之和，如两个原子的 AO（ψ_1，ψ_2）可组合成两个 MO（$\varPsi_1$，$\varPsi_2$），即

$$\varPsi_1 = C_1\psi_1 + C_2\psi_2 \text{，} \quad \varPsi_2 = C_1\psi_1 - C_2\psi_2 \tag{2-44}$$

式中，C_1，C_2 是线性组合系数（纯数）。组成的 MO 中通常一半是成键 MO，即两个原子轨道相加重叠而构成的分子轨道 $\varPsi_1$，一半是反键 MO，即两个原子轨道相减重叠而构成的分子轨道 $\varPsi_2$，有时还有未参与成键的为非键 MO。例如，第二周期中的某些元素的同核双原子分子如 O_2，F_2 等的 MO，可简单地由两个原子的 1s，2s，2p 的 10 个 AO 组合成 10 个 MO，其能量由低而高，大致为：σ_{1s}，σ_{1s}^*，σ_{2s}，σ_{2s}^*，σ_{2p_z}，(π_{2p_x}, π_{2p_y})，$(\pi_{2p_x}^*, \pi_{2p_y}^*)$，$\sigma_{2p_z}^*$（右上角*表示反键分子轨道）。它们的能级和电子云示意图如图 2-14 所示。

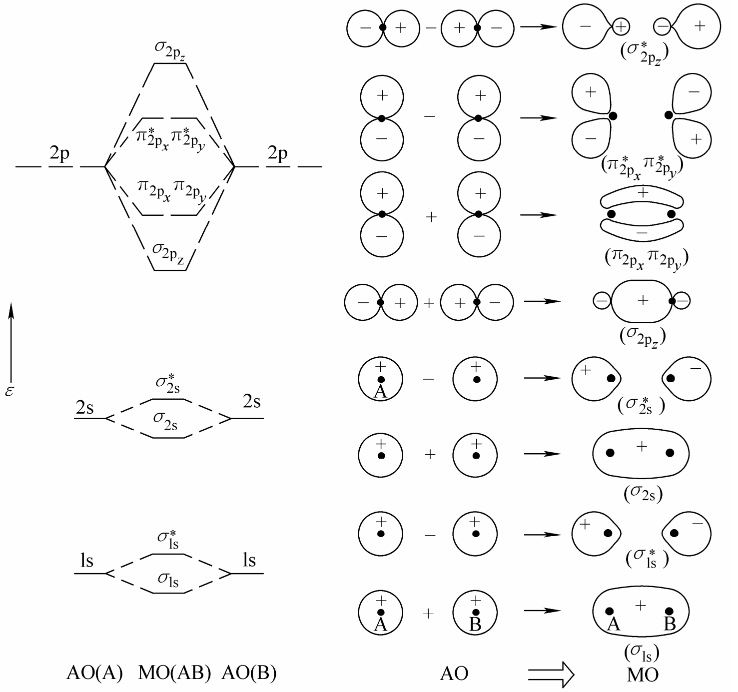


图 2-14 AO 构成 MO 及其电子云的示意图

3）MO 的 AO 有效组合原则　两个原子只能由能量相近的两个原子轨道才能有效地组合成两个 MO。两个原子轨道能量越相近，则所构成的成键 MO 越低于能量低的 AO，而反键 MO 则越高于能量高的 AO。若两个原子轨道能量相差悬殊，则不能有效组成 MO，电子基本上仍在各自 AO 上运动。当不同原子的价电子轨道能量相近，而有效地组合成分子轨道时，由于不同原子的不同价电子轨道能量各不相同，因此组合时其系数各不相同，能量也不同；凡参与成键的 AO 之间其波函数 ψ 符号相同部分重叠得越多者，成键的 MO 能量越低，越稳定；对称性匹配的 AO 相互重叠时方可使电子云达到最大重叠；泡利原理、能量最低原则及洪特规则同样适用于电子在 MO 上的填充和排布。

3. 金属键

周期表左面的元素均在满壳层外面有一个或多个价电子，这些填满的壳层在带正电的原子核和这些价电子之间起着屏蔽的作用，所以这些原子很容易失去它们的价电子而成为正离子。原子核的吸引力愈弱，则外面的电子“轨道”愈大。

在两个金属原子，例如两个钠原子互相接近的过程中，其（3s）“轨道”首先开始分裂。正如氢原子在相互接近时（1s）能级“轨道”发生分裂一样。如果这两个原子的（3s）“轨道”上的两个电子自旋相反，则结合成一个电子对，进入（3s）分裂后的较低能级的“轨道”，并使系统能量降低。假如又有一个也具有未成对的（3s）电子的钠原子接近这个电子对，则按照不相容原理，由于分裂后低能级的（3s）量子态已经填满，这个原子必然受到排斥。然而，实际上因为大量原子凝聚成固体时，最外面电子的能级展宽成带，钠的（3p）能级和（3s）能级发生重叠，所以第三个电子可以占据一个（3p）分裂后的低能级的量子状态以降低能量，当大量原子结合时，自旋相反的成对的价电子按不相容原理及能量最低原理排列在分裂展宽后（3s）及（3p）的较低能级的量子状态上，因而结合后系统总的能量下降。

上述金属原子分享价电子的设想，可导出金属键模型。在金属晶体中，金属原子上的价电子，在整个晶体内自由运动，金属晶体就是靠自由价电子和金属离子形成的点阵之间的相互作用结合在一起的。这种相互作用称为金属键。构成金属晶体的金属原子，由于价电子活动在整个金属晶体内，金属离子沉浸在运动的电子云雾中，所以电子云分布可以看成是球形对称的，整个原子可以看作具有一定体积的圆球，金属键没有饱和性和方向性。

金属的这个自由电子模型是德鲁特（Drude）在本世纪初提出的，它成功地对典型的金属作了定性的解释。金属自由电子的运动决定了金属的导电性，同时也决定了金属的其他一些特性。自由电子的运动以及原子本身的振动，使金属具有较大的热导率。金属内原子面之间作相对位移时，正离子仍和自由电子保持着结合键，因此金属具有可塑性。自由电子吸收可见光的能量使金属具有不透明性。

因吸收能量而激发的电子，当回到原来的低能级时，就发生辐射，这使金属具有光泽。

泡林（Pauling）在 20 世纪 30 年代提出了解释金属结合力的另一种观点。他认为金属结合是起源于相邻原子间价电子所形成的共价键。以碱金属钠为例，它具有体心立方点阵，每一个钠原子共有 14 个近邻（8 个近邻，6 个次近邻），中间的钠原子只有一个未成对的价电子，不足以在所有相邻原子间形成共价键，只能轮流和周围的原子形成单电子结合及双电子结合的共价键，所以可以把它看作是未饱和的共价键。泡林认为，如果结合键作无序的分布，每一个原子平均只参与一个结合键，而电子在这些键中共振，也使电子能量降低，即晶体的稳定性提高。所以，金属键被认为是一个共振过程，金属的不同原子之间可以形成一个电子或两个电子共价结合的各种结构，而金属键就是在这所有不同结构之间的电子的共振。

早期认为金属的传导性是由于电子在金属晶体中能自由移动的缘故。这种说法有时会造成错误印象，因为金刚石等共价键晶体具有交换效应，它们的电子也是从一个原子到另一个原子连续运动，整个电子群也被认为是在晶体中连续循环的，但金刚石等却不具有导电性。所以泡林等认为绝缘体和导体之间的不同，不在于电子是否被束缚或能自由移动，而是在外电场作用下，晶体中的电子能否在给定的方向流动。由于导体具有未被填满的电子状态（或电子“轨道”），在外加电场的作用下，电子可以经未填满的电子状态而通过。因此，凡是每个原子具有较少价电子的元素，具有金属导电性。因为这些元素凝聚后，可利用的状态没有被填满，共价键未饱和，电子在外加电场作用下，可通过未被填满的“轨道”而运动。

应用金属共振结合的观点，虽然能够解释不少实验事实，如金属的导电性、导热性、过渡族金属的强度和磁性等，但是由于共振结构只是在实验数据分析的基础上提出的假设，没有直接的实验证明，因此现代金属结合理论普遍应用的是较严格的金属能带理论。金属能带理论不但对金属的电学性质有较好的解释，而且它所作的假设也比较少。

4. 分子键

自由原子状态，已经形成稳定电子壳层的惰性气体 He、Ne、Ar 等和分子状态的 H_2、N_2、O_2 等在低温时都能结合成液态或固态。在它们结合成液态或固态的过程中，并不给出或接受电子。这种存在于中性的原子或分子之间的结合力称为分子键，也称范德华（Van der Waals）力，它是一种长程作用力。

图 2-15 所示为两个中性的原子，它们具有球形对称的电子云。它们互相接近时，若处于图 2-15a 所示瞬时状态，则形成偶极子，原子之间有一引力；若处于图 2-15b 所示瞬时状态，则原子之间有一斥力。按照量子力学的观点，非极性分

子（或原子）虽然本身偶极矩为零，但在运动的过程中，分子（或原子）上电子云分布的密度并不是始终均匀分布的，就每一个瞬间而言，由于电子云密度分布不均匀就产生了瞬时偶极矩，对靠近的非极性分子（或原子）的电子运动产生影响，形成偶极矩，因此发生相互作用而吸引。这就是范德华力。这种力很小，所以惰性气体的熔点和沸点都很低。

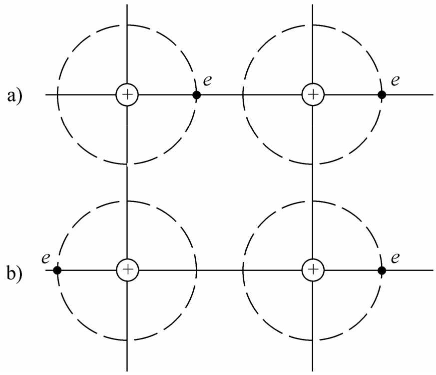


图 2-15　瞬时的电子状态

液态或固态的 H_2、N_2、O_2 也是靠范德华力进行凝聚的。它们的分子是靠共价键形成的。相邻的分子则由于电子互相协调运动而形成范德华力，一直到吸力和由于相邻分子的电子云重叠所产生的斥力相平衡为止。而石墨层的形成是由于其每个原子通过共价键与三个属于同一层的相邻原子相结合，每个原子的第四个电子形成金属键，为整个原子层所共有，但层间也是靠范德华力来结合的，所以层间结合较弱。

以上说明了在自由原子凝聚过程中，由于电子结构的变化，有四种不同的结合键及其相应的晶体。

在分子键晶体中，其原子的外层电子结构没有变化，如惰性气体。而其他三类结合键中，外层电子结构都发生了变化，有些原子给出了价电子。随着这些价电子被占有的方式的不同而分别形成三类不同的结合键晶体。在金属键晶体中，这些价电子为整个晶体所共有，例如单价金属，特别是碱金属；在共价键晶体中，这些价电子为邻近原子所共有，例如金刚石、硅、锗等；在离子键晶体中，这些价电子为其他原子所独占，例如碱金属和卤族元素形成 NaCl、LiF 等。事实上大多数晶体往往是以一种结合键为主，同时也综合有其他结合形式。

共价键晶体和离子键晶体的结合能约为 10^5（J/mol）数量级；金属键晶体的结合能约为 10^4（J/mol）数量级；分子键晶体的结合能约为 10^3（J/mol）数量级。因此，共价键晶体和离子键晶体结合最强，金属键晶体次之，分子键晶体最弱。

2.3　晶体结构

材料的力学性能和物理、化学性能是和构成材料的元素的原子结构、原子间的结合以及原子的排列、原子集合体的宏观形态有关。本节主要介绍晶体结构方面的知识。

2.3.1　晶体学基础

自然界存在的固态物质可分为晶体和非晶体两大类，晶体与非晶体的最本质的差别在于组成晶体的原子、离子、分子等质点是规则排列的，而非晶体中这些质点是无规则的堆积在一起的，非晶体的结构与液体结构相同，实际上非晶体就是过冷的、粘度很大的液体，或者说是被冻结的液体。

1. 晶体结构、空间点阵、晶胞、晶系

晶体的基本特征是原子（或分子、离子、原子集团）有规则的排列，现在来探讨由一种或多种原子所构成的晶体结构的种类。可能的晶体结构是无限多种的，但是可以利用晶体学方法对它进行分类。

为了研究晶体中物质质点排列的规律性，需要进行抽象，将实际存在的原子、离子或原子集团等物质质点抽象为纯粹的几何点，而完全忽略它的物质性，这样抽象出的几何点称为阵点或结点，阵点或结点在空间周期性的规则排列称为空间点阵。所以空间点阵是由具有物质性的原子、离子、原子集团的实际排列抽象出来的几何形象。空间点阵的主要特征是每个阵点或结点在空间都具有完全相同的环境，所谓完全相同的环境就是指在某个结点上向四周观察和从任何其他结点上向四周观察所见的景象完全相同。

必须强调指出，并不是晶体中的每一质点（原子、离子或原子集团）都必定和点阵中的结点相对应，例如 NaCl 的晶体就不属于简单立方点阵，而属于面心立方点阵。应该把每一对 Na^+与 Cl^-看成同属一结点。或者说，如果把 Na^+位置称为结点，则每一结点除有一个 Na^+之外还附有一个 Cl^-。空间点阵只有 14 种，而晶体结构却几乎是无限多种的。

在空间点阵中可以选择一个小的平行六面体作为基本单元，平行六面体的大小和形状可以用三维空间的 $\boldsymbol{a}$、$\boldsymbol{b}$、$\boldsymbol{c}$ 三个单位矢量来表示，这样的平行六面体称作晶胞。整个空间点阵可以看作是由很多大小和形状完全相同的晶胞紧密地堆垛在一起而形成的。矢量 $\boldsymbol{a}$、$\boldsymbol{b}$、$\boldsymbol{c}$ 不仅确定了单位晶胞，而且还确定了借这些矢量的平移而形成的整个点阵。即整个点阵中的所有阵点可以借矢量 $\boldsymbol{a}$、$\boldsymbol{b}$、$\boldsymbol{c}$ 由位于原点的一个阵点进行重复平移而产生。点阵中任何阵点的矢量坐标为 $P\boldsymbol{a}$、$Q\boldsymbol{b}$、$R\boldsymbol{c}$。对空间点阵的研究总是选取它的晶胞作为代表，通过对晶胞的研究就可以得知整

个空间点阵的情况。但是在同一空间点阵中有可能选择出多种不同形状和大小的平行六面体作为晶胞。究竟应如何选取晶胞呢？为了统一和方便，规定在选择晶胞时应满足下列条件：①晶胞的几何形状应与宏观晶体具有同样的对称性；②平行六面体内相等的棱和角的数目应最多；③当平行六面体的棱间存在直角时，直角的数目应该最多；④在满足上述条件的前提下晶胞应具有最小的体积。

晶胞的三个棱边分别用单位矢量 $\boldsymbol{a}$、$\boldsymbol{b}$、$\boldsymbol{c}$ 表示。$\boldsymbol{a}$、$\boldsymbol{b}$、$\boldsymbol{c}$ 之间相互夹角分别为 α（$\boldsymbol{c}$、$\boldsymbol{b}$ 间）、β（$\boldsymbol{a}$、$\boldsymbol{c}$ 间）、γ（$\boldsymbol{a}$、$\boldsymbol{b}$ 间）。a、b、c、α、β、γ 称为晶胞的六个参数，a、b、c 称为点阵常数。

法国晶体学家布喇菲（Bravais）曾用数学方法推算出所有晶体分属于 14 种空间点阵，根据晶胞的六个参数可以把这 14 种空间点阵归纳为七大晶系，如表 2-4 所示。14 种空间点阵如图 2-16 所示。在 14 种布喇菲晶胞中又可分为简单晶胞（亦称初基晶胞）和复合晶胞（亦称非初基晶胞），所谓简单晶胞即每个晶胞中只含有一个阵点，也就是只在晶胞的角顶位置上含有阵点。复合晶胞中则含有一个以上的阵点，除角顶位置外，在其他地方仍然有阵点，例体心，面心、底心等位置。每个晶胞的阵点数 N 可记为

$$N = N_{\rm i} + \frac{N_{\rm f}}{2} + \frac{N_{\rm c}}{8} \tag{2-45}$$

式中，$N_{\rm i}$ 为晶胞内的阵点数；$N_{\rm f}$ 为晶胞面上的阵点数；$N_{\rm c}$ 为晶胞角顶上的阵点数。

表 2-4　空间晶系与点阵

晶系	棱长及夹角关系	点阵类型	晶胞点阵数	点阵坐标
三斜晶系	$a\neq b\neq c, \alpha\neq\beta\neq\gamma\neq 90^\circ$	单简三斜	1	0 0 0
单斜晶系	$a\neq b\neq c, \alpha=\gamma=90^\circ\neq\beta$	单简单斜	1	0 0 0
		底心单斜	2	0 0 0, 1/2 1/2 0
六角晶系（六方）	$a=b\neq c, \alpha=\beta=90^\circ, \gamma=120^\circ$	简单六角	1	0 0 0
三角晶系（菱方）	$a=b=c, \alpha=\beta=\gamma\neq 90^\circ$	简单三角	1	0 0 0
正交晶系	$a\neq b\neq c, \alpha=\beta=\gamma=90^\circ$	简单正交	1	0 0 0
		体心正交	2	0 0 0, 1/2 1/2 1/2
		底心正交	2	0 0 0, 1/2 1/2 0
		面心正交	4	0 0 0, 1/2 1/2 0, 1/2 0 1/2, 0 1/2 1/2
四方晶系（正方）	$a=b\neq c, \alpha=\beta=\gamma=90^\circ$	简单四方	1	0 0 0
		体心四方	2	0 0 0, 1/2 1/2 1/2
立方晶系	$a=b=c, \alpha=\beta=\gamma=90^\circ$	简单立方	1	0 0 0
		体心立方	2	0 0 0, 1/2 1/2 1/2
		面心立方	4	0 0 0, 1/2 1/2 0, 1/2 0 1/2, 0 1/2 1/2

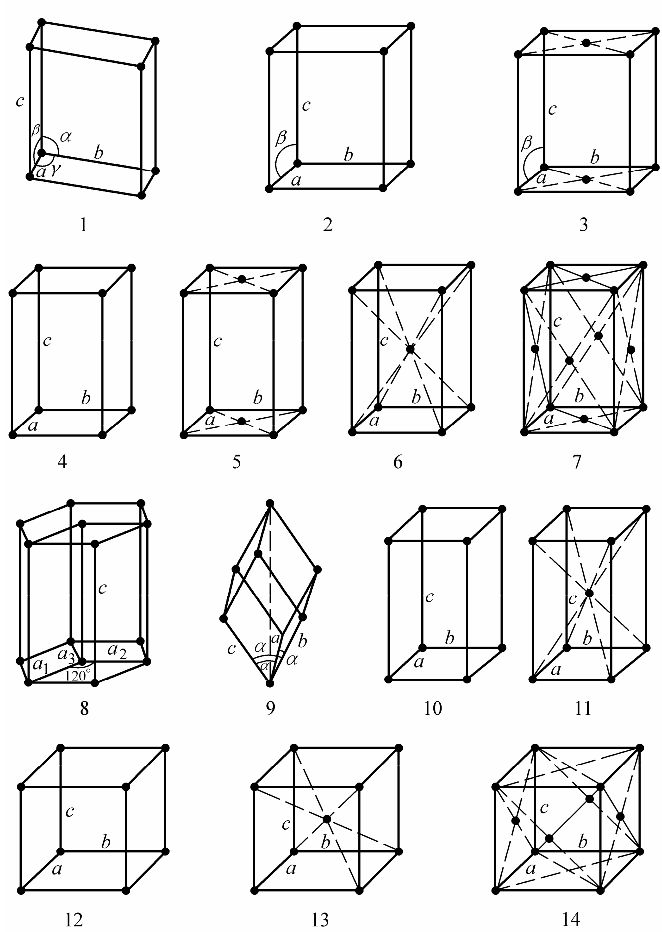


图 2-16　14 种布喇菲空间点阵的晶胞

1—简单三斜　2—简单单斜　3—底心单斜　4—简单正交　5—底心正交　6—体心正交　7—面心正交

8—简单六方　9—菱形（三角）　10—简单四方　11—体心四方　12—简单立方　13—体心立方　14—面心立方

这 14 种点阵外的其他任何形式的点阵都可以用这 14 种点阵之一来表示。例如底心四方点阵可以用简单四方表示，如图 2-17 所示。面心四方点阵可以用体心四方来表示，如图 2-18 所示。

前面讲到在点阵中任何阵点的矢量坐标可以记为 $P\boldsymbol{a}$、$Q\boldsymbol{b}$、$R\boldsymbol{c}$，当采用简单晶胞时，P、Q、R 都是整数，而当采用复合晶胞时 P、Q、R 可能是分数。虽然复合晶胞也可以改用简单晶胞来表示，例如面心立方晶胞可以用简单三角晶胞来表示，但采用复合的面心立方晶胞更合适，因为面心立方晶胞具有高度对称性。在 14 种布喇菲点阵中共有 8 种复合晶胞。

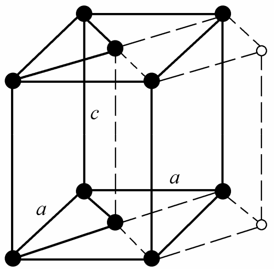


图 2-17　底心四方点阵和简单四方点阵的关系

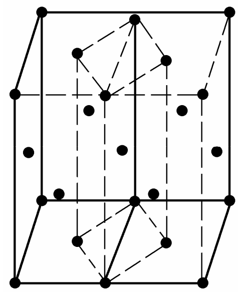


图 2-18　面心四方点阵和体心四方点阵的关系

晶体虽然一共只有 14 种布喇菲点阵，但由于每个阵点可以存在着一个或多个同种或异种物质质点（原子、离子、原子集团），而且这些质点在阵点上的排列组合图案也可以有多种形式，所以每种空间点阵可以统括无限多种结构的晶体。

2. 晶面指数、晶向指数

在晶体中由物质质点所组成的平面称为晶面，物质质点所组成的直线称为晶向。在晶体中各不同晶向和晶面上原子（或离子、原子集团，下同）排列的密度不一样，这种结构上的差异引起晶体在各个方向上的物理、化学、力学性能上存在差异，这就是晶体的各向异性。在各种物理化学过程中（例如弹性变形、塑性变形、相变等），不同的晶向和晶面的表现往往存在着差异，而某些特殊的晶面（例如最密排面）往往起着特殊的作用，为了区别和研究的方便，有必要用适当的指数符号来标定晶体中的晶面和晶向。

（1）晶面指数标定

在晶体中，原子的排列构成了许多不同的晶面，故要用晶面指数来分别表示这些晶面。晶面指数的确定方法如下：①在正交晶系中，在点阵中设立参考坐标系，使坐标系的三个轴与晶胞的三个棱边重合，以晶胞的点阵常数 a、b、c 分别为三个坐标轴的单位长度；②求出被标定晶面与 x、y、z 三个坐标轴相交所得截距，为了避免出现零截距，故所选坐标系原点一定要在被标定晶面之外，而不能在被标定晶面上；③取各截距倒数；④将三个倒数化为最小的简单整数，并加上圆括号，即表示该晶面的指数，一般记为(hkl)。图 2-19 所示的 ABC 晶面的指数为（236）。如果被标定晶面与坐标轴的负方向相截，则在指数数字上方冠以负号。

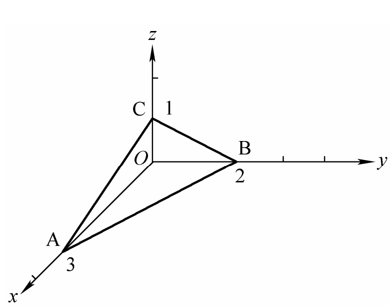


图 2-19　ABC 晶面指数的标定

在晶体中凡位于坐标系的同一象限中互相平行的平面都具有同一晶面指数。如果在晶面指数上都冠以负号，则所代表晶面仍与未冠负号的晶面互相平行，只是不在坐标系的同一象限中。所以晶面指数(hkl)不是指一个晶面，而是代表相互平行的一组晶面。在晶体中有些晶面具有共同的特点，其上原子排列和分布是完全相同的，晶面间距也相同，只是晶面的位向不同，这样的几个等同晶面称为一个晶面族。在立方系中这样的晶面族中的各个晶面的指数数字相同，但数字的排列次序和正负号不同。晶面族符号记为$\{hkl\}$。在立方系中$\{hk1\}$晶面族中所包括的晶面可以用h、k、l数字的排列组合方法求之。例如{100}包括(100)、(010)、(001)。{110}包括(110)、(101)、(011)、$(\bar{1}10)$、$(\bar{1}01)$、$(0\bar{1}1)$。{111}包括 (111)、$(\bar{1}11)$、$(1\bar{1}1)$、$(11\bar{1})$。$\{hkl\}$晶面族中所包含的晶面数可这样计算：若h、k、l中三个数都不相等，且都不等0，则此晶面族包括有3!×4组晶面，3!为h、k、l的排列数；如果在h、k、l中有两个相等则包括有(3!/2!)×4组晶面；如果在h、k、l中有一个0存在，则包括有(3!/2)×4组晶面。

（2）晶向指数的标定

假设有一组互相平行的晶向L_0、L_1 、L_2…，其中L_0通过原点O，如图2-20所示。设R_0为L_0上最接近原点的结点，其坐标为(x_0, y_0, z_0)，则$\boldsymbol{OR}_0$为L_0上单位矢量。如果$\boldsymbol{a}$、$\boldsymbol{b}$、$\boldsymbol{c}$为过原点O的三个坐标轴上的单位矢量，则$\boldsymbol{OR}_0 = x_0\boldsymbol{a}+ y_0\boldsymbol{b} +z_0\boldsymbol{c}$。设$R_1(x_1, y_1, z_1)$为$L_1$上一点，而$R\,(x, y, z)$为$L_1$上任一点，则$\boldsymbol{OR}_1 = x_1\boldsymbol{a}+ y_1\boldsymbol{b} +z_1\boldsymbol{c}$，$\boldsymbol{OR} = x\boldsymbol{a}+ y\boldsymbol{b} +z\boldsymbol{c}$，而$\boldsymbol{RR}_1 = \boldsymbol{OR} - \boldsymbol{OR}_1$。由于同一方向上（即相互平行晶向上）结点是周期性重复的，所以$\boldsymbol{RR}_1 = \boldsymbol{OR} - \boldsymbol{OR}_1 = n(\boldsymbol{OR}_0)$，$n$为整数。故

$$(x-x_1)\boldsymbol{a}+(y-y_1)\boldsymbol{b}+(z-z_1)\boldsymbol{c}=n(x_0\boldsymbol{a}+y_0\boldsymbol{b}+z_0\boldsymbol{c}) \tag{2-46}$$

比较式（2-46）两边$\boldsymbol{a}$、$\boldsymbol{b}$、$\boldsymbol{c}$的系数，得到L_1的方程式为

$$\frac{x-x_1}{x_0}=\frac{y-y_1}{y_0}=\frac{z-z_1}{z_0} \tag{2-47}$$

对于L_2、L_3等其他晶向同样存在有方程式

$$L_2: \frac{x-x_2}{x_0}=\frac{y-y_2}{y_0}=\frac{z-z_2}{z_0} \tag{2-48}$$

$$L_3: \frac{x-x_3}{x_0}=\frac{y-y_3}{y_0}=\frac{z-z_3}{z_0} \tag{2-49}$$

由此得出

$$\begin{aligned} x_0:y_0:z_0 &= (x-x_1):(y-y_1):(z-z_1) \\ &= (x-x_2):(y-y_2):(z-z_2) \\ &= (x-x_3):(y-y_3):(z-z_3) \end{aligned} \tag{2-50}$$

所以利用$x_0 : y_0 : z_0$可以表示一束平行的晶向，它还可以化为三个互质数u、v 、w

的连比，即

$$x_0 : y_0 : z_0 = u : v : w \tag{2-51}$$

则 u、v、w 为晶向 L_0、L_1、L_2…的晶向指数，记为[$u\,v\,w$]。所以晶向指数的求法有两种：① 在坐标系中过原点做被标定晶向的平行线 L_0，取 L_0 上任一结点坐标 (x, y, z)，将其化为三个互质整数的连比 $x: y: z= u: v: w$，则[$u\,v\,w$]为晶向指数。② 在坐标系中写出晶向 L 上任意二结点的坐标(x_1, y_1, z_1)和(x_2, y_2, z_2)，将(x_1-x_2): (y_1-y_2): (z_1-z_2)化为三个互质整数之连比即可得到晶向指数。同样，在晶体结构中那些原子密度相同的等同晶向同属一个晶向族，用符号<$u\,v\,w$>表示。

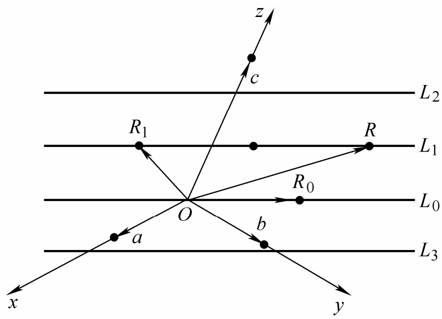


图 2-20　推导晶向方程式的示意图

凡平行于同一直线的晶面都属于一个晶带，而该直线则称为晶带轴。同一晶带的晶面其晶面指数和晶面间距可能完全不同，但它们必须都与晶带轴平行。当晶带轴的指数为[$u\,v\,w$]时，则晶带中任何一个晶面的指数(hkl)都必须满足 $hu+kv+lw=0$。任何两个非平行的晶面都属于同一晶带，因为它们一定都平行于其交线，如果此两晶面指数分别为$(h_1k_1l_1)$和$(h_2k_2\,l_2)$，则其交线即晶带轴的指数为

$$\left.\begin{aligned} u &= k_1l_2 - k_2l_1 \\ v &= l_1h_2 - l_2h_1 \\ w &= h_1k_2 - h_2k_1 \end{aligned}\right\} \tag{2-52}$$

如果有两晶向$[u_1\,v_1\,w_1]$和$[u_2\,v_2\,w_2]$，则此两晶向所决定的晶面指数为

$$h:k:l = \begin{vmatrix} v_1 & w_1 \\ v_2 & w_2 \end{vmatrix} : \begin{vmatrix} w_1 & u_1 \\ w_2 & u_2 \end{vmatrix} : \begin{vmatrix} u_1 & v_1 \\ u_2 & v_2 \end{vmatrix} \tag{2-53}$$

在立方晶系中指数相同的晶面和晶向是互相垂直的，例如[110]⊥(110)和[111]⊥(111)。晶体结构中的各组晶面，其面间距的数值各不相同，面间距大的晶面，其指数比较低，而面间距小的晶面指数比较高。在立方系中晶面间距可由下式求出（a 为晶胞棱长）

$$d_{hkl}=\frac{a}{\sqrt{h^2+k^2+l^2}} \tag{2-54}$$

如果两个晶向的指数分别为$[u_1\,v_1\,w_1]$和$[u_2\,v_2\,w_2]$，当 $\alpha=\beta=\gamma=90°$时（正交晶系，α、β、γ 晶胞棱的夹角），两个晶向间所夹角 Φ 的余弦（a、b、c 为晶胞棱长）

$$\cos\Phi=\frac{a^2u_1u_2+b^2v_1v_2+c^2w_1w_2}{\sqrt{a^2u_1^2+b^2v_1^2+c^2w_1^2}\sqrt{a^2u_2^2+b^2v_2^2+c^2w_2^2}} \tag{2-55}$$

如果一晶面指数为(hkl)，一晶向指数为$[u\,v\,w]$，则晶面与晶向间夹角 Φ 可以认为是晶向与晶面法线间夹角的余角，即 $\Phi=90°-\psi$，其中 ψ 为晶向与晶面法线间夹角，当 $\alpha=\beta=\gamma=90°$时，有

$$\cos\psi=\sin\Phi=\frac{hu+kv+lw}{\sqrt{\frac{h^2}{a^2}+\frac{k^2}{b^2}+\frac{l^2}{c^2}}\sqrt{a^2u^2+b^2v^2+c^2w^2}} \tag{2-56}$$

（3）六方系晶面晶向指数的标定

对于六方晶系也可以采用三轴坐标系，底面上的 x_1、x_2 轴之间夹角为 120°，z 轴垂直于 Ox_1x_2 平面。在此坐标系中，确定晶面指数的方法仍按前所述，六方柱体的六个侧面的晶面指数分别为(100)、(010)、$(\bar{1}10)$、$(\bar{1}00)$、$(0\bar{1}0)$、$(1\bar{1}0)$。六方柱体的各个侧面上原子排列相同，且所在位置也保持一定的对称关系，属于等同面，但从它们的晶面指数上却不能明确表示出来。在此三轴坐标系中晶向指数标定方法亦与前相同。

为了显示六方晶系的对称性，常常采用四轴坐标系，即在 Ox_1x_2 平面上，加一轴 Ox_3，且使其与 Ox_1、Ox_2 均成 120°，如图 2-21 所示。此时晶面指数的标定方法不变，则六方柱体的六个侧面的晶面指数分别为$(10\bar{1}0)$、$(01\bar{1}0)$、$(\bar{1}100)$、$(\bar{1}010)$、$(0\bar{1}10)$、$(1\bar{1}00)$，它可以显示出六方晶面的等同性和对称性，它的一般形式记为$(hkil)$，可以证明 $i=-(h+k)$。在四轴坐标系中，由于点的坐标与三轴坐标有所不同，故晶向指数的标定方法也不同，四轴坐标系中晶向指数一般记为$[uvtw]$。三轴坐标系与四轴坐标系之间存在一定的换算关系：假设有一点 R，在三轴坐标系中坐标为(x, y, z)，其平面坐标为(x, y)；在四轴坐标系中坐标为(x_1, x_2, x_3, z)，其平面坐标为(x_1, x_2, x_3)，但在平面上用三个坐标值来表示一个点，可以有多种不定解，故此点 R 还必须满足

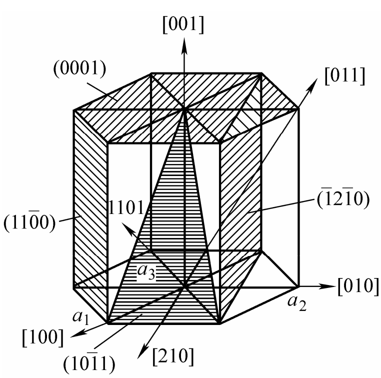


图 2-21 六方晶系晶面晶向指数

$$x_1 + x_2 + x_3 = 0 \tag{2-57}$$

坐标系中单位矢量要满足

$$\boldsymbol{a}_1 + \boldsymbol{a}_2 + \boldsymbol{a}_3 = 0 \tag{2-58}$$

因此，

$$\boldsymbol{OR} = x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + x_3\boldsymbol{a}_3 + z\boldsymbol{c} = (x_1 - x_3)\boldsymbol{a}_1 + (x_2 - x_3)\boldsymbol{a}_2 + z\boldsymbol{c} \tag{2-59}$$

假设 $a = a_1$、$b = a_2$，则根据

$$\boldsymbol{OR} = x\boldsymbol{a} + y\boldsymbol{b} + z\boldsymbol{c} \tag{2-60}$$

可以得到

$$\left.\begin{aligned} x &= x_1 - x_3 \\ y &= x_2 - x_3 \\ z &= z \end{aligned}\right\} \tag{2-61}$$

$$\left.\begin{aligned} x_1 &= \frac{2}{3}x - \frac{1}{3}y \\ x_2 &= \frac{2}{3}y - \frac{1}{3}x \\ x_3 &= -\frac{1}{3}x - \frac{1}{3}y \\ z &= z \end{aligned}\right\} \tag{2-62}$$

式（2-61）和式（2-62）为 x、y、z 三轴坐标系与 x_1、x_2、x_3、z 四轴坐标系点的坐标的换算关系式。

对于晶向指数三轴坐标系[UVW]与四轴坐标系[$uvtw$]亦有转换关系式，由于 $u+v+t = 0$，故

$$\left.\begin{aligned} U &= u - t \\ V &= v - t \\ W &= w \end{aligned}\right\} \tag{2-63}$$

$$\left.\begin{aligned} u &= \frac{2}{3}U - \frac{1}{3}V \\ v &= \frac{2}{3}V - \frac{1}{3}U \\ t &= -\frac{1}{3}U - \frac{1}{3}V \\ w &= W \end{aligned}\right\} \tag{2-64}$$

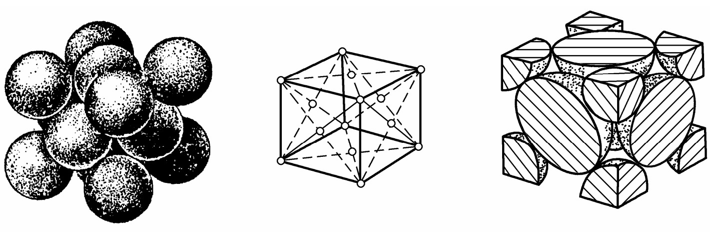
2.3.2　典型晶体结构

几乎所有元素的晶体结构都被人们测定过，其中较简单的三种典型的晶体结

构为面心立方、体心立方和密排六方。

1. **面心立方**

面心立方晶胞如图 2-22 所示，在晶胞中原子位于六面体的八个顶点和六个表面的面心位置，即(0, 0, 0)，(1/2, 1/2, 0)，(1/2, 0, 1/2)，(0, 1/2, 1/2)。根据式（2-45），每个晶胞内包含有四个原子。晶胞的点阵常数可以用 a 来表示。可以将原子看作是半径相等的，相互吸引的小钢球，由于能量条件的要求，这些小钢球总是趋于最紧密的排列，最邻近的小钢球彼此相切，相切的两个原子中心之间的距离为原子直径。在面心立方中最密排面是{111}，密排方向是<110>，原子直径为 $(\sqrt{2}/2)a$，原子半径则为 $(\sqrt{2}/4)a$。在{111}面上每一列原子都是紧密相切的，第二列紧密排列的原子位于第一列原子的空隙处，而第三列原子又位于第二列原子的空隙处，依次排列构成{111}密排面，如图 2-23 所示。在密排面上每个原子和周围 6 个原子相切，假设在密排面内各小钢球的位置用 A 来表示，则间隙位置有两种，分别以 B 和 C 表示。第二层排面上的原子排在 B 位置，第三层密排面上的原子位于 C 位置，一直到第四层密排面原子才重复出现在 A 位置，如图 2-23 所示。所以面心立方密排面的堆垛次序为 $ABCABC\cdots$。



a) b) c)

图 2-22 面心立方晶胞

a）钢球模型 b）质点模型 c）晶胞中原子数示意图

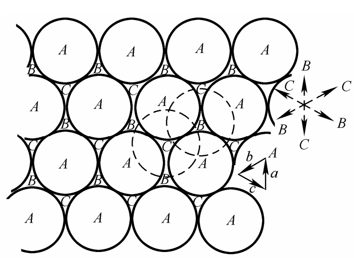


图 2-23 面心立方{111}面上原子排列及{111}面原子堆垛次序

在晶体中每一原子周围所具有的，与其等距离的最近邻的原子数目叫做配位数。很显然，在面心立方晶体中，和某一原子相距为 $(\sqrt{2}/2)a$ 的最近邻原子数有

12 个，故配位数为 12。

原子钢球在空间堆垛的紧密程度可以用致密度来衡量，致密度亦称为堆跺密度。致密度以单位晶胞中原子钢球所占体积和晶胞的体积之比来表示。在面心立方晶胞中，原子所占体积为 74%，而余下的 26%则为晶胞内的空隙体积。面心立方晶胞中的间隙有两种，一种称为八面体间隙，间隙的中心位于晶胞的体心位置和晶胞每个棱边的中点，它是由周围六个原子所围成的八面体中心，故称为八面体间隙。在面心立方晶胞中，八面体间隙的半径为原子半径的 0.41 倍；另一种间隙称为四面体间隙，它是由一个顶角原子和相邻的三个面心原子所组成的四面体所围成。四面体间隙的半径为原子半径的 0.225 倍。如图 2-24 和图 2-25 所示。在八面体间隙和四面体间隙处常常可以容纳某些半径比较小的杂质原子或溶质原子。

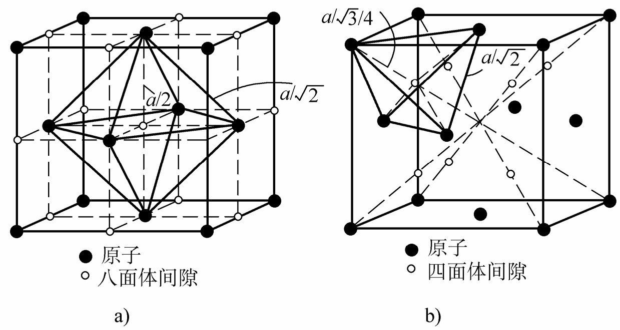


图 2-24　面心立方结构的间隙

a）八面体间隙　b）四面体间隙

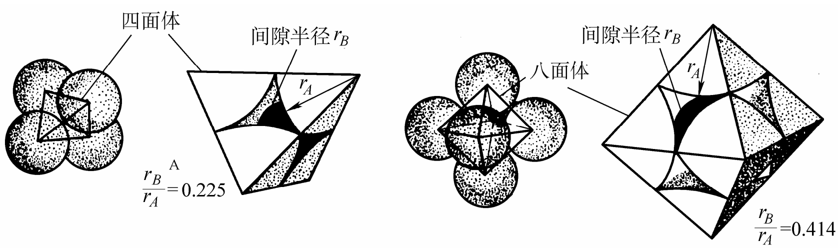


图 2-25　面心立方的八面体间隙和四面体间隙的钢球模型

2. 体心立方

体心立方晶胞如图 2-26 所示。原子位于六面体的每个顶角和体心位置，即原子坐标为(0, 0, 0)和(1/2, 1/2, 1/2)，每个晶胞中有 2 个原子，晶胞的点阵常数可以用 a 来表示。最密排方向为体对角线方向，即<111>。在<111>方向上原子紧密排列，彼此相切，原子半径为 $(\sqrt{3}/4)\,a$，每一个原子周围有 8 个最近邻原子，因此配位

数为 8。但每个原子周围除有 8 个最近邻原子外，尚有 6 个次近邻原子，次近邻原子的距离比最近邻原子距离远约 15%左右，因此往往要考虑到次近邻原子的作用，有时将配位数记为 8+6，即有效配位数要大于 8。在体心立方晶胞中，没有类似面心立方{111}的那种最密排面，原子排列比较紧密的面是{110}，其上有<111>族的两个最密排方向。体心立方晶胞的致密度为 0.68。它的致密度比面心立方小，总间隙比较大。它的间隙位置也分为两种，如图 2-27 所示。一种是扁八面体间隙，位于晶胞六面体的每个面的中心位置和每个棱的中心，八面体间隙半径为原子半径的 0.154 倍。另一种是四面体间隙位于由两个体心原子和两个顶角原子所组成的四面体中心，四面体间隙半径为原子半径的 0.291 倍。由此看出虽然体心立方的总间隙体积比面心立方的大，但从它的间隙尺寸分析则每个间隙体积比较小，而且它的四面体间隙尺寸比八面体间隙要大，这和面心立方情况相反，因此在体心立方晶体的间隙中所能容纳的杂质原子或溶质原子量往往比面心立方要少。

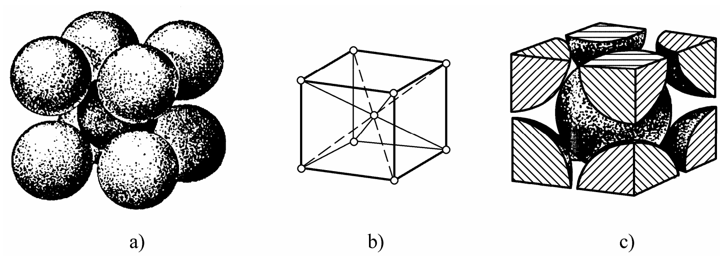


图 2-26 体心立方晶胞

a）钢球模型 b）质点模型 c）晶胞中原子数示意图

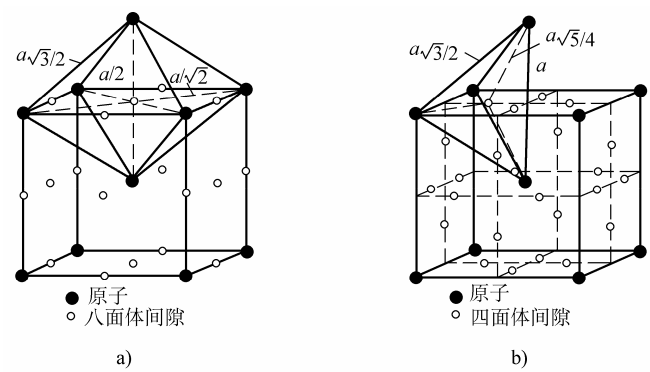


图 2-27 体心立方结构的间隙

a）八面体间隙 b）四面体间隙

3. 密排六方结构

密排六方的晶胞可以有两种取法，一种取法为一平行六面体，每个晶胞中包

含有位于晶胞内和八个顶点位置的原子共 2 个。但为了更好地反映晶体的对称性而选取六棱柱体为晶胞，如图 2-28 所示，此晶胞的体积恰好为平行六面体的 3 倍，这样的晶胞中包含有 6 个原子，晶胞内有 3 个，12 个顶角原子和上下两个底心原子。晶格常数用 a、c 两个数值表示，a 表示底面的六角形边长，c 表示棱柱体的高度。晶胞内的原子位于 $c=1/2$ 处，它们在底面上的投影分别落在三个相间的三角形的中心位置。在理想情况下（最密排情况），底面上原子之间距离和上下层间原子间距相等，此时轴比 $c/a=\sqrt{8/3}=1.633$。但实际上很多呈这种结构的晶体其轴比在 1.633 附近变动。当同一底面上原子之间距离大于上下层的原子之间距离时，$c/a<1.633$；当底面上原子之间距离小于上下层的原子之间距离时，轴比 $c/a>1.633$。当 $c/a=1.633$ 时，原子半径为 $a/2$，配位数为 12，致密度约为 0.74。当轴比偏离 1.633 时，配位数可记为 6 +6。在密排六方结构中，最密排面是(0001)，其上有 3 个最密排方向$[11\bar{2}0]$，$[\bar{2}110]$和$[1\bar{2}10]$。在密排面(0001)上原子排列情况和面心立方{111}上原子排列情况完全一样，只是密排面的堆垛次序不一样。在面心立方中堆垛次序为 *ABCABC*…，而在密排六方中堆垛次序为 *ABABAB*…。密排六方结构中总的间隙体积与面心立方相同，它的间隙位置也分为两种，如图 2-29 所

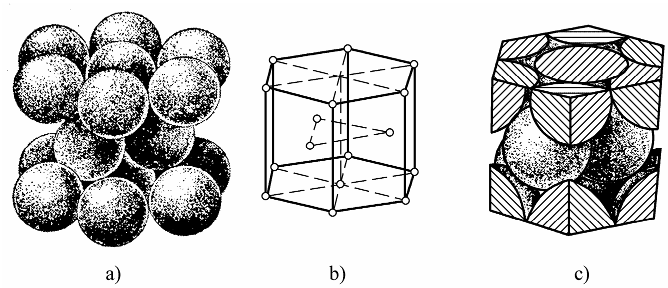


图 2-28　密排六方结构晶胞

a）钢球模型　b）质点模型　c）晶胞中原子数示意图

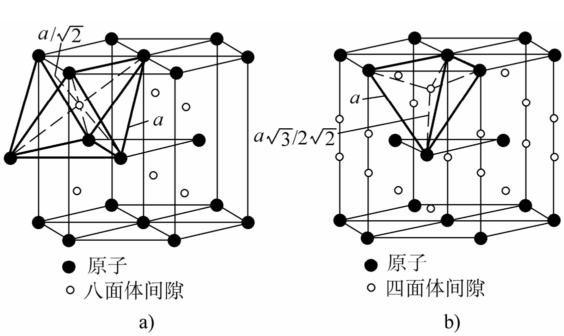


图 2-29　密排六方结构的间隙

a）八面体间隙　b）四面体间隙

示，一种为八面体间隙，间隙半径为原子半径的 0.41 倍，另一种为四面体间隙，间隙半径为原子半径的 0.225 倍。最后要指出的是密排六方结构是属于简单六方点阵的，因为晶胞内处于(0, 0, 0)和(2/3, 1/3, 1/2)处的二类原子的环境是不同的，不能同时作为阵点，如果顶角为阵点，那么体内三个原子应分别附属于一个顶角，这就是说每个阵点包含着两个原子，所以密排六方结构属于简单六方点阵。

4. 其他类型结构

（1）复杂密排结构

在密排结构中，除了密排次序为 *ABCABC*…的面心立方和 *ABAB*…的密排六方结构以外，尚有更复杂的密排结构。已测定出一种四层密排结构，堆踩次序为〔*ABAC*〕*ABAC*…，在镧系元素中的 α-镨、α-钕等和锕系中的 α-镅都具有这种结构。另有一种九层的密排结构，堆踩次序为〔*ABCBCACAB*〕…，镧系元素中的钐具有这种结构。在各种密排结构中往往由于加工或相变等造成堆垛次序的错误，因而形成错层。

（2）金刚石结构

金刚石结构如图 2-30 所示，这种晶体结构也属于面心立方点阵。每一阵点上有两个原子。可以看成是两个面心立方点阵穿插而成。两个面心立方点阵沿着体对角线相对位移了体对角线的 1/4 长度。每个晶胞中包含有 8 个原子，原子之间最近距离为 $(\sqrt{3}/4)\,a$。每一个原子被 4 个邻近原子所包围，此四个邻近原子组成一个四面体，该原子位于四面体中心，这种结构的配位为 4。致密度约为 0.344。

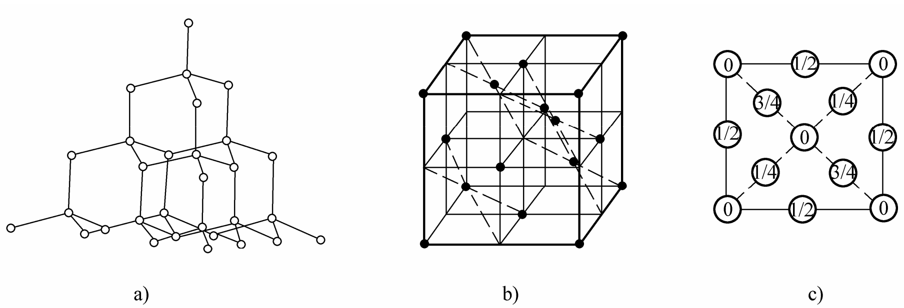


图 2-30 金刚石结构

a）共价键 b）晶胞 c）原子在底面的投影

（3）其他亚金属结构

第 VA 族元素 As、Sb、Bi 具有菱方结构，如图 2-31 所示。配位数等于 3，每个原子有三个最近邻的原子，以共价键方式相结合并形成层状结构（图 2-31 中水平面）。层与层之间结合带有金属键的性质。所以，这几种亚金属兼有金属和非金属的特性。图 2-32 第 VIB 族元素 Se 和 Te 的晶体结构，其配位数等于 2，每个

原子有两个近邻，以共价键方式相结合。原子组成呈螺旋分布的链状结构。其结构可归为菱方，当采用六方系坐标轴时，用 c 和 a 两个点阵常数来表明晶胞的大小。螺旋链绕着 c 轴，原子与 c 轴距离为 x，Se 和 Te 的 x 值分别等于 $0.217a$ 和 $0.269a$。

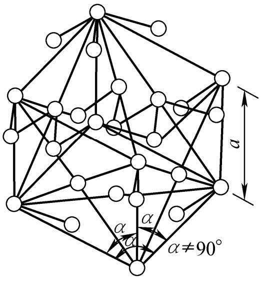


图 2-31　第 VA 族元素 As、Sb、Bi 的晶体结构

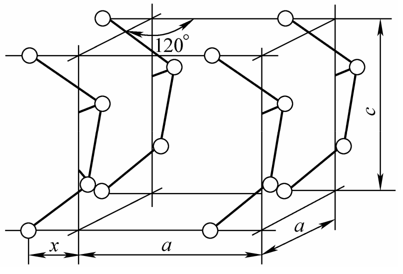


图 2-32　Se 和 Te 的晶体结构

第 3 章　金属电子论

金属电子论亦称金属态理论，它从单个原子结构出发研究金属原子集合体中电子（主要是原子最外层活跃的价电子）的能量状态，并用它来解释金属及合金的结构与性质，包括金属材料的电、磁、热、弹等物理性能。早期电子论主要涉及金属的导电性及导热性，后来发展成为解释金属态及一切固体许多重要特性的理论基础。电子理论的发展经历了三个阶段，即经典自由电子论、量子自由电子论和能带论。经典自由电子论认为金属中的价电子是完全自由的，并服从经典力学规律。量子自由电子论也认为金属中价电子完全自由，但服从量子力学规律。能带论则考虑到点阵周期场的作用，并用量子力学观点处理问题。本章概要介绍金属电子论的一些基本物理概念及有关情况，以便为以后各章提供必要的理论基础。

3.1　经典自由电子论

3.1.1　经典自由电子论的基本假设

原子中不是所有电子在决定固体金属的强度、硬度等特性的相互作用中都是起同样作用的。X 光的谱线明显地说明，形成原子满壳层的内部电子实质上是没有影响的，只是外面的价电子在金属键中起主要的作用。德鲁特—洛伦兹（Drude-Lorents）的经典自由电子论对上述这两类电子作了严格的区分，主要考虑了价电子的作用，即想象在一块金属中，所有原子的价电子是共有的，价电子组成一种能在金属整体中自由运动的“电子气”。其运动遵守经典力学定律，特别是遵守理想气体分子运动规律。忽略价电子间的排斥作用，以及正离子所造成的点阵周期场对价电子的作用。

3.1.2　经典自由电子论的主要成就

这个理论在定性地解释金属的一些主要特性，如导电、导热等方面取得了满意的结果。

1. 经典自由电子论的欧姆定律推导

欧姆定律原是电学中根据实验概括出来的经验规律，经典自由电子论从理论导出了这个规律。在一金属导体的长度方向施加电场 $\boldsymbol{E}$，电子在电场作用下的加

速度为 $\boldsymbol{a}$，除无规则的热运动外电子在电场作用下还产生一个附加的定向运动。电子在正离子组成的晶格中运动，电子与正离子发生碰撞。电子与正离子每次碰撞后的一瞬间，电子在电场作用下的定向速度可以看作零。设电子热运动平均速度大小为 $\overline{v}_t$，两次碰撞的平均自由程为 $\overline{L}$，自由飞行时间为 $\overline{\tau}$。实验表明，施加电场后电子的定向速度比热运动速度小许多倍，故在计算 $\overline{\tau}$ 时，定向运动速度忽略不计，即 $\overline{\tau}=\overline{L}/\overline{\upsilon}_t$。碰撞前一瞬间的定向速度为 ν_d（$=\boldsymbol{a}\overline{\tau}$），那么，在两次碰撞间电子定向平均速度 $\overline{v}_d$ 为

$$\overline{v}_d=\frac{1}{2}\nu_d=\frac{1}{2}\boldsymbol{a}\overline{\tau}=\frac{1}{2}\frac{e\boldsymbol{E}}{m}\overline{\tau}=\frac{e\overline{L}}{2m\overline{v}_t}\boldsymbol{E} \tag{3-1}$$

设导体单位体积内的电子数为 n，则单位时间内通过单位面积内的电子数为 $n\overline{v}_d$（单位时间电子的飞行距离即为 $\overline{v}_d$ 值的大小），单位时间内通过此单位面积的电荷为 $ne\overline{v}_d$，故电流密度 δ 大小为

$$\boldsymbol{\delta}=ne\overline{v}_d=\frac{ne^2\overline{L}}{2m\overline{v}_t}E \tag{3-2}$$

对于一定的金属导体，在一定的温度下，$\dfrac{ne^2\overline{L}}{2m\overline{v}_t}$ 是常数，故式（3-2）表示电流密度与电场强度成正比，这就是欧姆定律。式（3-2）还包括另一条定律，叫导电定律，其表达式为

$$\sigma=\frac{ne^2}{2m}\frac{\overline{L}}{\overline{v}_t}=\frac{ne^2}{2m}\overline{\tau} \tag{3-3}$$

式中，σ 为电导率。

2. 经典自由电子论的焦耳—楞次定律推导

电子与离子碰撞时，电子把定向运动的那部分动能传递给正离子，结果是离子热振动加剧，导体温度上升。当点阵离子所获得的能量与热传导所散失的能量平衡时，导体温度即不再上升。每次碰撞，电子定向运动的动能变为热能 ΔW，根据上述，ΔW 为

$$\Delta W=\frac{1}{2}m{\nu_d}^2=\frac{1}{2}\frac{e^2E^2}{m}\overline{\tau}^2 \tag{3-4}$$

单位时间内每一个电子和点阵离子平均碰撞 $1/\overline{\tau}$ 次数，故单位时间内电子总共传递给单位体积金属的热能 W 为

$$W=n\frac{1}{\overline{\tau}}\Delta W=\frac{ne^2}{2m}\overline{\tau}E^2=\sigma E^2 \tag{3-5}$$

式（3-5）即为焦耳—楞次定律。

经典电子理论还可以导出维德曼—佛朗兹定律，成功地解释金属具有良好的导电性及导热性等。这些成就的取得绝非偶然，因为经典电子理论中的一些假设

基本上是对的，例如经典电子理论认为自由电子能够在整个金属中运动，提出这一点就等于找到了理解金属态的钥匙。

3.1.3　经典自由电子论遇到的困难及原因

经典自由电子论导出的电子热容以及导电率与温度的关系与实验结果不符。

1. 电子热容问题

根据经典自由电子论，金属中的价电子如同理想气体，在温度 T 下每个电子的平均能量为(3/2)kT，k 为玻耳兹曼常数。对一价金属而言，每摩尔原子的电子气的能量 $E_e = N_A(3/2)kT = (3/2)RT$，式中 N_A 为阿伏加德罗常数，R 为摩尔气体常数。一摩尔原子的电子气摩尔热容 $C_{V,\mathrm{m}}^{e} = \dfrac{\mathrm{d}E_e}{\mathrm{d}T} = \dfrac{3}{2}R$ =12.471J/(K·mol)。这一结果比实验数据高出 100 倍。

2. 电导率与温度的关系

在导电定律式（3-3）的诸参量中，只有电子热运动的平均速度 $\bar{\upsilon}_t$ 与温度 T 有关，按理想气体分子热运动定律，$(1/2)m\bar{v}_t^2 = (3/2)kT$，那么有

$$\sigma = \frac{e^2 n\bar{L}}{2(3mkT)^{1/2}} \tag{3-6}$$

式（3-6）表明电导率 σ 与 $T^{1/2}$ 成反比，但实验表明 σ 与 T 成反比。

经典电子理论另一个不能解释的问题是关于一、二价金属的导电问题。按式（3-3），σ 与 n 成正比，由此可得出二价金属的导电性应比一价金属好，但实际情况并非如此。

经典电子论遇到了不可克服的困难。其根源在于经典电子论立足于牛顿力学，而对微观粒子运动问题，需用量子力学解决。

3.2　量子自由电子理论

量子自由电子论的基本假设是：金属离子所形成的势场各处是均匀的；价电子是共有的，可以在整个金属内自由运动，价电子之间以及价电子与离子之间没有相互作用；电子运动服从量子力学原理。量子自由电子理论与经典电子理论的主要区别在于电子运动服从量子力学原理。

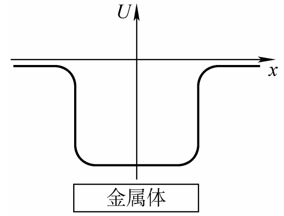


图 3-1　电子在金属中的势能曲线

电子在保守力场的作用下在金属中运动。由于电子要逸出金属需克服正电荷的吸引，因此电子在金属外的电势能高于金属内的电势能，其一维的势能曲线如图 3-1 所示，其形状与陷阱相似，

故称为势阱。

3.2.1　阱内势能为零的电子状态

1. 一维势阱

（1）一维无限深势阱

一个理想的势阱模型—无限深势阱。阱内（金属内）电子完全自由，只有动能，没有势能；阱外势能无限高，即

$$U(x)=\begin{cases}0 & 0<x<L \quad（阱内）\\ \infty & x\leq 0,\ x\geq L（阱外）\end{cases} \tag{3-7}$$

其势能曲线如图 3-2 所示。

质量为 m、能量为 ε 的自由电子被封闭在如图 3-2 所示的长度为 L 的一维势阱中。描述阱内电子运动的薛定谔方程为

$$\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2}+\frac{2m}{\hbar^2}\varepsilon\psi=0 \tag{3-8}$$

式中 $\hbar$ 为约化普朗克常数。令

$$k=\sqrt{\frac{2m\varepsilon}{\hbar^2}}=\frac{2\pi}{\lambda} \tag{3-9}$$

这里 k 就是波数。式（3-8）可改写为

$$\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2}+k^2\psi=0 \tag{3-10}$$

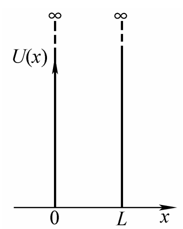


图 3-2　一维无限深势阱

方程式（3-10）的解为

$$\psi(x)=A\sin(kx+\delta) \tag{3-11}$$

因是无限深势阱，有边界条件 $\psi(0)=0$ 和 $\psi(L)=0$。将边界条件代入式（3-11）则有 $\delta=0$ 和 $\sin(kL)=0$，从而得到

$$kL=n\pi \quad (n=1,2,3\cdots) \tag{3-12a}$$

式（3-12a）中 $n=0$ 时，$\varepsilon=0$，舍去此状态。式（3-11）中的常数 A 可根据 0～L 区间的波函数归一化条件求出，归一化条件为

$$\int_0^L\left|\psi(x)\right|^2\mathrm{d}x=1 \tag{3-12b}$$

求出的 A 值为$(2/L)^{1/2}$。将 δ 和 A 值以及式（3-12b）代入式（3-10）得

$$\psi(x)=\sqrt{\frac{2}{L}}\sin\left(\frac{n\pi}{L}x\right) \tag{3-13}$$

将式（3-12a）代入式（3-9）得

$$\varepsilon=\frac{\hbar^2k^2}{2m}=\frac{\hbar^2}{2m}\left(\frac{n\pi}{L}\right)^2=\varepsilon_n \quad (n=1,2,3\cdots) \tag{3-14}$$

式（3-14）表明金属中的电子所有的能量是量子化的，有正整数 n 确定，n 叫量子数。图 3-3 是处于势阱中电子的波函数及波函数的平方。由图 3-3 可见，在不同部位找到电子的几率是不同的。

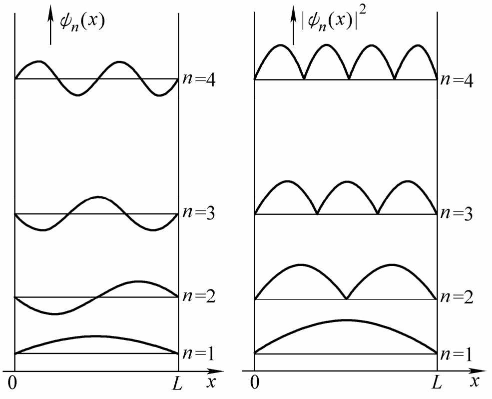


图 3-3　处于势阱中电子的波函数及波函数的平方

（2）一维有限深势阱

实际金属表面势垒的高度是有限的（宽度也较小），所以在 $x = 0$ 和 $x = L$ 处的边界条件不容易确定，为此，假定波函数 $\psi(x)$具有长度为 L 的周期，如图 3-4a 所示。于是，边界条件为

$$\psi(x) = \psi(x + L) \tag{3-15}$$

这个边界条件相当于把长度为 L 的一维势场两端连接成如图 3-4b 所示的环形，并假定波函数 $\psi(x)$具有长度为 L 的周期。这个假定完全是一种数学技巧，其目的是避开具体假定边界条件的困难。由于固体内部的性质不受表面的影响，因此这样处理是允许的。对波函数假定的条件式（3-15）称为周期性边界条件。

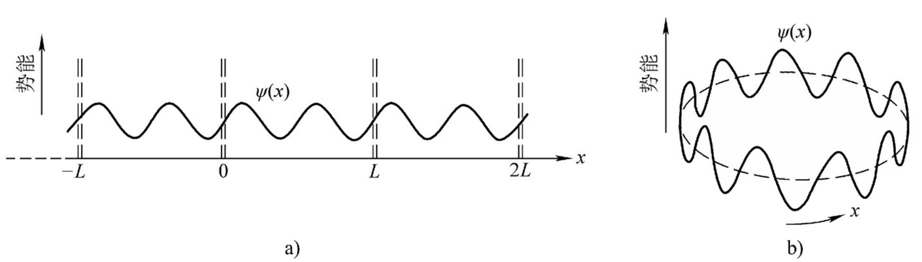


图 3-4　周期性边界条件

a）周期波函数　b）环形周期波函数

势阱内的波函数和解也分别是式（3-10）和式（3-11）。将周期性边界条件式（3-15）代入式（3-11）可得 $kL= 2n\pi$，n 为正整数（$n = 0$ 舍去），即

$$k = \pm\frac{2\pi}{L},\ \pm\frac{4\pi}{L},\ \pm\frac{6\pi}{L},\ \pm\frac{8\pi}{L},\ \cdots \tag{3-16}$$

将式（3-16）代入式（3-9）得

$$\varepsilon = \frac{\hbar^2 k^2}{2m} = \frac{\hbar^2}{2m}\left(\frac{2n\pi}{L}\right)^2 = \varepsilon_n \quad (n = 1, 2, 3\cdots) \tag{3-17}$$

由式（3-17）可以看出，ε 与 k 为抛物线关系，在 k 轴上，每个能量状态所占长度为 $2\pi/L$。波函数中的 δ 和 A 值无法求得。

2. 三维有限势阱

上面讨论了一维有限深势阱，现在讨论封闭在边长为 L 的立方势箱（有限势阱）中质量为 m、势能为 ε 的电子运动状态。设波矢为 $\boldsymbol{k}$，方向与电子运动方向相同，$k = 2\pi/\lambda = (2m\varepsilon/\hbar^2)^{1/2}$。波动方程为

$$\nabla^2\psi_k(\boldsymbol{r}) + \frac{2m}{\hbar^2}\varepsilon_k\psi_k(\boldsymbol{r}) = 0 \tag{3-18}$$

其波函数为

$$\psi_k(r) = A\exp[\mathrm{i}(kr + \delta)] = A\exp[\mathrm{i}(k_x x + k_y y + k_z z + \delta)] \tag{3-19}$$

波函数和能量的脚标 k 表示状态是由波矢 $\boldsymbol{k}$ 决定的。

和一维的情况一样，假设 x、y、z 各方向具有周期性的边界条件为

$$\left.\begin{aligned}\psi_k(x+L, y, z) = \psi_k(x, y, z)\\ \psi_k(x, y+L, z) = \psi_k(x, y, z)\\ \psi_k(x, y, z+L) = \psi_k(x, y, z)\end{aligned}\right\} \tag{3-20}$$

式中 $k^2 = k_x^2 + k_y^2 + k_z^2$。将式（3-20）代入式（3-19）得所允许的波矢 $\boldsymbol{k}$ 的三个分量为

$$k_x, k_y, k_z = \pm\frac{2\pi}{L},\ \pm\frac{4\pi}{L},\ \pm\frac{6\pi}{L},\ \pm\frac{8\pi}{L},\cdots \tag{3-21}$$

本征能量值 ε_k 为

$$\varepsilon_k = \frac{\hbar^2 k^2}{2m} = \frac{\hbar^2}{2m}(k_x^2 + k_y^2 + k_z^2) \tag{3-22}$$

对于 k_x、k_y、k_z，每隔 $2\pi/L$ 有一个量子状态，波矢 $\boldsymbol{k}$ 的三个分量的值决定后，能量状态 ε_k 就决定了。

3.2.2　三维自由电子

1. 费米半径和费米能

边长为 L 的立方体晶体块，其内部的电子可以看成三维有限势阱内的电子。根据式（3-22），在波矢 $\boldsymbol{k}$ 空间（以 k_x、k_y、k_z 为坐标建立的波数空间）中，等能面

（ε_k 为常数）为球面，如图 3-5 所示。k_x、k_y 和 k_z 均可以取正值、负值和零。电子的能量与 k^2 成正比。由式（3-22）可知，具有相同的 $\boldsymbol{k}^2$ 值的点具有相同的能量，例如（$2\pi/L$,0,0）、（0,$2\pi/L$,0）和（0,0,$2\pi/L$）坐标点的能量值相等，却表示不同的量子状态（量子状态不同于能量状态），电子由靠近原点低能量状态依次填充到更高的能量状态。设有 N 个电子占据了从原点到半径 $k = k_F$ 的球面（称为费米球）之间的能量状态。$\boldsymbol{k}$ 空间放大时，$\boldsymbol{k}$ 空间中每一个微小体积$(2\pi/L)^3$ 就有一个量子状态，考虑到电子的自旋，一个量子状态可以容纳两个电子。半径为 k_F 的球的体积为$(4/3)\pi k_F{}^3$，因此有

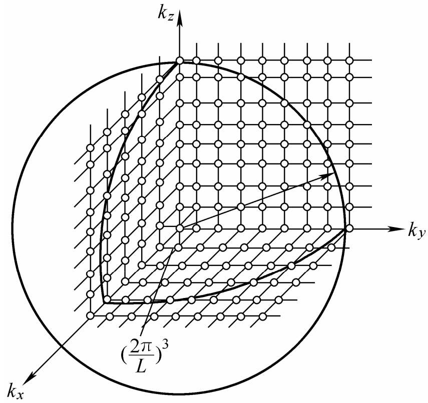


图 3-5　三维自由电子的能量状态和费米球

$$2\frac{(4/3)\pi k_F^3}{(2\pi/L)^3}=\frac{L^3}{3\pi^2}k_F^3=N \tag{3-23}$$

因此费米球半径 k_F 为

$$k_F=\left(\frac{3\pi^2 N}{V}\right)^{1/3} \tag{3-24}$$

式中体积 $V=L^3$。费米面上的电子速度（费米速度）v_F 为

$$v_F=\frac{\hbar k_F}{m}=\frac{\hbar}{m}\left(\frac{3\pi^2 N}{V}\right)^{1/3} \tag{3-25}$$

根据式（3-22），费米球表面处的能量（费米能）ε_F 为

$$\varepsilon_F=\frac{\hbar^2 k_F{}^2}{2m}=\frac{\hbar^2}{2m}\left(\frac{3\pi^2 N}{V}\right)^{2/3} \tag{3-26}$$

普通金属中 ε_F 约为 5eV，v_F 大约为 1.5×10^{16}m/s。

2. 费米—狄拉克统计和靠近费米能级的能量

电子数的能量分布情况可用状态密度来描述。状态密度就是晶体内单位能量间隔的状态数（即电子数），用 $\rho(\varepsilon)$表示。如图 3-6 所示，在 $\boldsymbol{k}$ 空间中，由能量 ε 到 ε+dε 的两个等能面所包围区域的体积为 $4\pi k^2dk$，每一个微小体积$(2\pi/L)^3$ 就有一个量子状态，所以考虑到自旋并利用式（3-22），得到三维情况下的状态（不是能量状态）密度函数 $\rho(\varepsilon)$为

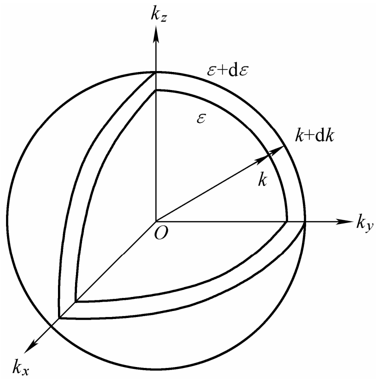


图 3-6　三维自由电子的等能面

$$\rho(\varepsilon)\mathrm{d}\varepsilon = 2\frac{4\pi k^2 \mathrm{d}k}{(2\pi/L)^3} = \frac{L^3}{2\pi^2}\left(\frac{2m}{\hbar^2}\right)^{3/2}\varepsilon^{1/2}\mathrm{d}\varepsilon \tag{3-27}$$

即

$$\rho(\varepsilon) = 2\frac{4\pi k^2 \mathrm{d}k}{(2\pi/L)^3} = \frac{L^3}{2\pi^2}\left(\frac{2m}{\hbar^2}\right)^{3/2}\varepsilon^{1/2} \tag{3-28}$$

$\rho(\varepsilon)$与 ε 呈抛物线关系。在 T= 0K 时，电子最大的能量为 $\varepsilon_F(0)$（0K 时的费米能），在 $\varepsilon_F(0)$以上，各态均无电子，在 $\varepsilon_F(0)$以下各态均填满电子，而且愈接近 $\varepsilon_F(0)$其状态密度愈大，如图 3-7a 所示。

在任意温度 T（单位为 K）时，热平衡状态下，电子处于能量为 ε 的几率，可用费米—狄拉克分布函数描述。由于电子比一般气体分子轻很多，金属中电子密度比标准条件下一般气体分子高 10^4 倍，电子的能量变化是量子化的，因此对金属中的电子，经典力学范畴的麦克斯韦—玻耳兹曼统计已不合适，需用量子统计，或称费米—狄拉克分布（或统计）。该统计的基本假设是：电子彼此不能区分；一个能带的每一个能级最多只能被两个自旋相反的电子占据。该统计的关系式为

$$f(\varepsilon) = \frac{1}{\exp[(\varepsilon - \varepsilon_F)/kT] + 1} \tag{3-29}$$

式中 $f(\varepsilon)$为费米—狄拉克分布函数，它表示在热平衡时，电子处于能量为 ε 的几率；ε_F 为温度为 T 时的费米能；这里的 k 为玻尔兹曼常数。分布函数 f 具有下列性质：当 T= 0 K 时，若 $\varepsilon > \varepsilon_F(0)$，$f$ = 0；若 $\varepsilon < \varepsilon_F(0)$，$f$ = 1，如图 3-7b 实线所示。就是说为了使系统总能量最小，电子倾向于占有能量低的状态，能量小于 $\varepsilon_F(0)$的状态均被电子占有；能量大于 $\varepsilon_F(0)$的状态全部空着。当 T > 0K，f 的分布如图 3-7b 中虚线所示，若 $\varepsilon = \varepsilon_F$ ，f=1/2（在图 3-7b 中，视 $\varepsilon_F \approx \varepsilon_F(0)$）。在室温下 $kT \approx 0.025$eV，而金属的 ε_F 为几电子伏特。因此只有少量和能量 ε_F 接近的电子能吸收热能而跃迁

到能量较高的状态，故高于 ε_F 的原来空的能态也有部分被电子占据。能量比 ε_F 低更多的电子不能吸收热能，因为 kT 能量不足以把这些电子激发到更高能级，而相近的能级则已被电子填满。这是量子力学处理电子分布问题所得出的崭新的概念。

可以导出在 0 K 时电子的平均动能为$(3/5)\varepsilon_F(0)$，说明在绝对零度，电子仍然具有相当大的平均动能。这也是与经典理论截然不同的概念。

当 $T > 0\text{K}$，且 $kT < \varepsilon_F(F)$时，有

$$\varepsilon_F(F) = \varepsilon_F(0)\left[1 - \frac{\pi^2}{12}\left(\frac{kT}{\varepsilon_F(0)}\right)^2\right] \tag{3-30}$$

由式（3-30）可见，当温度升高时，$\varepsilon_F(F)$要从 $\varepsilon_F(0)$下降。但因 $kT \ll \varepsilon_F(0)$，所以直到金属处于熔化状态的温度时，ε_F 与 $\varepsilon_F(0)$相差很小，故可以把费米能看作一个常数。

在温度不为零时，电子的状态密度分布如图 3-7c 所示。在 ε_F 周围大约 $2kT$ 的范围内，电子占据情况变成虚线的样子。

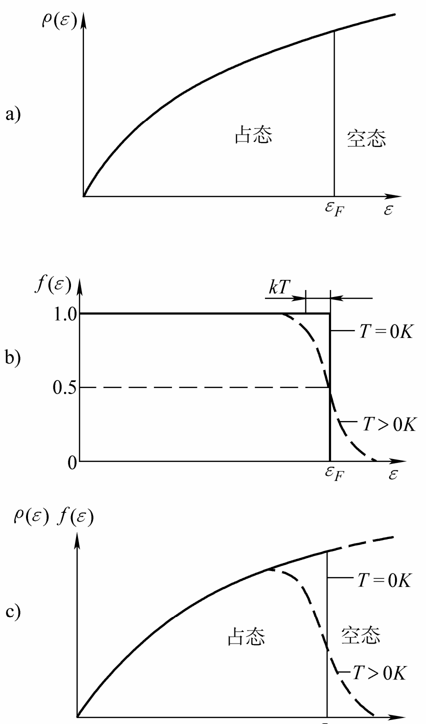


图 3-7 靠近费米能级时电子状态分配特点

a）0K 时状态密度曲线 b）电子状态几率分布曲线 c）几率分布对电子状态分配的影响

3.2.3　量子自由电子理论对金属导电及电子摩尔热容的解释及遇到的问题

在前面介绍经典自由电子论时，曾提到经典自由电子论在解释金属导电及电子摩尔比热容问题时遇到的困难。这些困难按量子自由电子论可以解决。

1. 电子摩尔比热容

在经典理论中，电子能量分布服从麦克斯韦—玻尔兹曼统计，每个自由度具有(3/2)kT的热能，总数为 N 的自由电子的摩尔比热容为(3/2)Nk。然而，金属电子摩尔比热容的实测值约为上述数值的 1/100。这种不相符合的原因在于，金属中自由电子的分布服从费米—狄喇克统计，因此对摩尔比热容有贡献的只限于费米面附近，宽度约为 kT 的范围中存在的很少一部分电子。

金属的电子摩尔比热容可以用下面的方法大致估算。在温度作用下电子要从外部吸收热能跃迁到另一能级，根据泡利不相容原理，后一能级必须是空能级，只有在费米能量 ε_F 附近大约 kT 能量范围内的电子满足这个条件，若总电子数为 N，这部分电子的数量约为

$$N\frac{kT}{\varepsilon_F}=N\frac{T}{T_F} \tag{3-31}$$

式中，T_F 为由 $\varepsilon_F=kT_F$ 决定的费米温度。这些电子平均获得大约 kT 的热能，所以电子系统的总热能 Q 为

$$Q=kTN\frac{T}{T_F} \tag{3-32}$$

电子数为 N 的电子摩尔热容 C_{eN} 为

$$C_{eN}=\frac{\mathrm{d}Q}{\mathrm{d}T}=2Nk\frac{T}{T_F} \tag{3-33}$$

可以看出，金属的电子摩尔比热容比经典的电子理论导出的结果(3/2)NkT 小一个因子 T/T_F，金属的电子摩尔比热容和温度成正比。利用费米—狄拉克分布进行精确计算，金属的电子摩尔比热容为

$$C_{eN}=\frac{\pi^2}{2}Nk\frac{T}{T_F} \tag{3-34}$$

对于普通金属，由上式计算出的摩尔热容约为 10^{-4}J/（mol·K）。在室温下，这个值约为晶格振动摩尔比热容的 1/100（也是经典电子理论导出的电子摩尔比热容的 1/100）。因此，电子对于金属摩尔比热容的贡献很小。但是，在低温时晶格摩尔比热容随温度的下降与 T^3 成比例地减小，所以在低温时，电子摩尔比热容的贡献就变得很主要了。

2. 金属导电

量子自由电子论用费米—狄拉克分布的观点解释金属导电问题。图 3-8 是费米—狄拉克分布描述一维金属导电的示意图。图中纵坐标 f 表示状态被电子占据

的几率，横坐标 K 为各状态下的波数，K_m 是与 $\varepsilon_f(0)$ 相应的波数。在 0K 下，在 $\varepsilon_f(0)$ 以下的状态都被电子填满，相应地在图 3-8 中 $+K_m$ 与 $-K_m$ 之间的状态几率都相等（$f=1$）。在未加电场时，虽然电子在金属中运动，但是如有一个电子以某一速度沿 $+K$ 方向运动，则必有另一个电子以相同的速度向相反方向运动。在未加电场前，在 $+K_m$ 与 $-K_m$ 间电子状态是对原点对称分布的，在 $+K_m$ 与 $-K_m$ 区间之外的状态都还空着（$f=1$），所以不呈现电流。施加电场之后，电子沿 $+K$ 方向受到加速，此时整个分布曲线向右偏移，如图 3-8 虚线所示。在此情况下，大多数电子仍然不产生电流，因为它们的 $+K$ 与 $-K$ 方向相互抵消。只有那些波数接近 $+K_m$ 的电子由于没有相应的 $-K_m$ 电子与之抵消，所以它们沿 $+K$ 方向运动而产生电流。可见，虽然所有的电子都受到电场的加速，但只有那些在费米—狄拉克分布曲线端部的电子能参加导电，而经典电子论则认为所有价电子都参加导电，这是两者的区别。

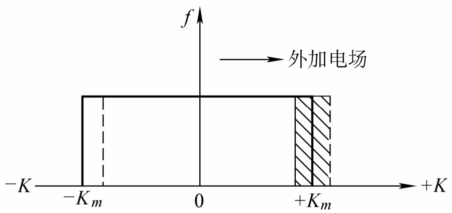


图 3-8　用费米—狄拉克分布描述导电示意图

原子结合成为晶体时，有的成为导体，有的成为半导体，有的成为绝缘体，而量子自由电子理论不能解释这些问题。

3.3　周期势场和布洛赫（A. Bloch）定律

3.3.1　周期势场

金属通过金属键结合成晶格，金属中的价电子在经典电子论和量子自由电子论中均认为是在晶体中自由运动的电子。但是，在实际金属中，电子既受到图 3-9 所示的由规则排列的原子实正电荷的负库仑引力势场和其他价电子的斥力势场迭加而成的周期势场 V 的作用，在逸出时又受到表面势垒的作用，要使电子逸出金属，必须补给电子如图 3-9 所示的能量 Φ_W（功函数）。

原子结合成为晶体时，有的成为导体，有的成为绝缘体，其原因只能用周期势场的作用来说明。在研究规则排列的晶格原子周期势场的作用时，为简单起见，在不影响问题本质的范围内作如下的假定：①晶格原子静止在平衡位置，而晶格振动作为散射电子的微扰来考虑；②传导电子不但通过库仑斥力直接相互作用，而且通过对晶格原子势场的屏蔽效应间接地相互影响，因此严格地说，这是个多

电子间问题。但是，可以近似地把每个电子看成是在由晶格原子形成的周期势场和其他电子分布形成的平均势场迭加而成的周期势场中相互独立地运动着。这种近似称为单电子近似。

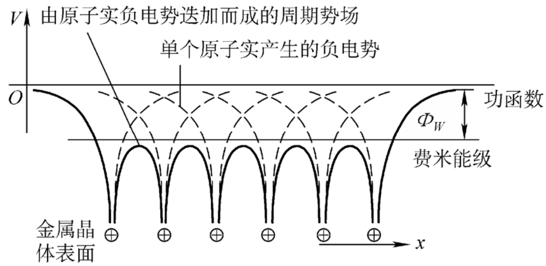


图 3-9　金属中的周期势场

3.3.2　布洛赫定律

如图 3-10a 所示，以周期 a 排列的 N 个原子构成长度为 L 的一维晶体周期势场 $V(x)$，在 $V(x)$作用下的电子波动方程为

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2\psi(x)}{\mathrm{d}x^2}+V(x)\psi(x)=\varepsilon\psi(x) \tag{3-35}$$

因为势场 $V(x)$和晶体具有相同的周期，所以

$$V(x+a)=V(x) \tag{3-36}$$

把一维晶体左右无限地连接起来或者把它的两端连接成环行，引入波函数的周期性边界条件

$$\psi(x+L)=\psi(x+Na)=\psi(x) \tag{3-37}$$

将式（3-35）中的 x 变换为 $x+a$，则有

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2\psi(x+a)}{\mathrm{d}x^2}+V(x+a)\psi(x+a)=\varepsilon\psi(x+a) \tag{3-38}$$

考虑到式（3-36）的条件，把式（3-38）和式（3-35）相比较，可以看出 $\psi(x)$和 $\psi(x+a)$是属于同一波动方程的同一本征值的两个本征函数，因此如果能量状态不简并（简并时也能证明得到下面的同样结果），则 $\psi(x)$和 $\psi(x+a)$线性相关，它们之间只差一个系数，即

$$\psi(x+a)=\lambda\psi(x) \tag{3-39}$$

对坐标 x 重复进行 N 次长度为 a 的平移操作，则由式（3-39）得

$$\psi(x+L)=\psi(x+Na)=\lambda^N\psi(x) \tag{3-40}$$

根据式（3-37）和式（3-40），则 $\lambda^N=1$，即 λ 为 1 的 N 次方根。λ 可表示为

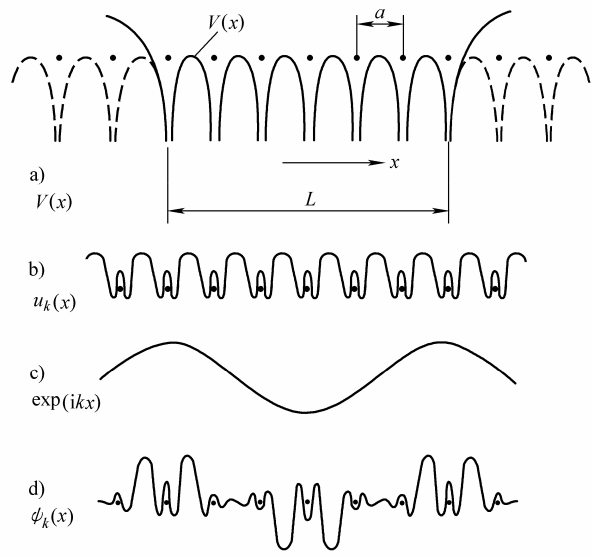


图 3-10 布洛赫函数的周期性

a）一维晶体的周期势场 b）与晶格有相同周期函数 $u_k(x)$

c）自由电子的平面波 $\exp(ikx)$ d）布洛赫函数 $\psi_k(x)$

$$\lambda = \exp\left(\frac{\mathrm{i}2\pi g}{N}\right) \tag{3-41}$$

式中，g 是 $0 \leqslant g \leqslant N-1$ 区间中的整数。

因此式（3-39）可写为

$$\psi(x+a) = \exp\left(\frac{\mathrm{i}2\pi g}{N}\right)\psi(x) \tag{3-42}$$

将式（3-42）两边同乘 $\exp[-\mathrm{i}2\pi(x+a)/(Na)]$得

$$\exp\left[-\frac{\mathrm{i}2\pi g(x+a)}{Na}\right]\psi(x+a) = \exp\left(-\frac{\mathrm{i}2\pi gx}{Na}\right)\psi(x) \tag{3-43}$$

令

$$u_g(x) = \exp\left(-\frac{\mathrm{i}2\pi gx}{Na}\right)\psi(x) \tag{3-44}$$

由式（3-43）可以看出，$u_g(x)$是具有和晶格势场有相同周期 a 的周期函数，如图 3-10b 所示。根据式（3-44），在周期势场作用下的运动电子波函数 $\psi(x)$可表示为

$$\psi_k(x) = u_g(x)\exp\left(\frac{\mathrm{i}2\pi gx}{Na}\right) = u_k(x)\exp(\mathrm{i}kx) \tag{3-45}$$

式中波数 $k = 2\pi g/(Na)$。式（3-45）称为布洛赫函数，布洛赫函数 $\psi_k(x)$是由与周期

势场有相同周期的函数 $u_k(x)$调制的平面波 $\exp(\mathrm{i}kx)$（自由电子的波函数），如图 3-10d 所示。式（3-45）中脚标 k 表示波函数 $\psi(x)$是由分立的波数 k 确定的。布洛赫函数 $\psi_k(x)$具有下面的性质

$$\psi_k(x+a)=\exp(\mathrm{i}ka)u_k(x) \tag{3-46}$$

从波数 k 的定义 $k=2\pi g/(Na)=2\pi g/L$ 可以看出，k 许可的值是分立的，和一维势阱中的自由电子的情况一样，这些值在 k 轴上以 $2\pi/L$ 的间隔排列着。

把一维布洛赫函数推广到三维时，采用波矢 $\boldsymbol{k}(k_x, k_y, k_z)$和位置矢量 $\boldsymbol{r}(x, y, z)$，则三维布洛赫函数为

$$\psi_k(\boldsymbol{r})=u_k(\boldsymbol{r})\exp(\mathrm{i}\boldsymbol{kr}) \tag{3-47}$$

布洛赫证明了一条重要定理，即对于含周期势的薛定谔方程，其解必定具有布洛赫函数的形式。

3.3.3　克勒尼希—彭尼（Kronig-Penney）模型

如上所述，周期势场中的电子波函数是布洛赫函数。然而，一般对于式（3-35）无法精确求解。下面采用一种近似方法。这里介绍克勒尼希—彭尼的非常简单的一维模型，由这个模型能够对式（3-35）求出解析形式的解，并能定性地说明周期势场中电子运动情况的本质。

如图 3-11 所示，将一维晶体的晶格势场 $V(x)$近似成宽度为 b，周期为 $a+b$（后面设 $b\to0$，这时周期即为 a），高度为 V_0 的方势阱。波动方程和波函数的布洛赫函数为

$$\frac{\mathrm{d}^2\psi(x)}{\mathrm{d}x^2}+\frac{2m}{\hbar^2}\left[\varepsilon-V(x)\right]\psi(x)=0 \tag{3-48}$$

$$\psi(x)=u(x)\exp(\mathrm{i}kx) \tag{3-49}$$

式中

$$\left.\begin{array}{ll} V(x)=V_0 & (-b\leqslant x<0) \\ V(x)=0 & (0\leqslant x<a) \end{array}\right\} \tag{3-50}$$

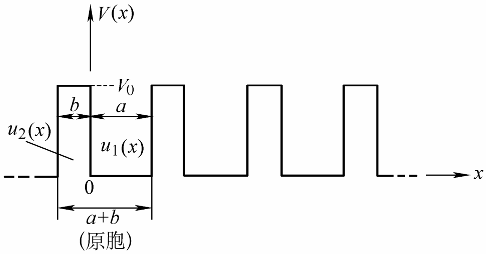


图 3-11　克勒尼希—彭尼模型中势阱

$u(x)$是以 $a+b$ 为周期的周期函数，波数 $k = 2\pi g/(Na)$，g 是 $0\leqslant g\leqslant N-1$ 区间中的整数。

将式（3-49）代入式（3-48）得

$$\frac{\mathrm{d}^2u(x)}{\mathrm{d}x^2}+2\mathrm{i}k\frac{\mathrm{d}u(x)}{\mathrm{d}x}-\left(k^2-\alpha^2+\frac{2mV(x)}{\hbar^2}\right)u(x)=0 \qquad (3\text{-}51)$$

式中

$$\alpha=\left(\frac{2m\varepsilon}{\hbar^2}\right)^{1/2} \qquad (3\text{-}52)$$

设式(3-49)中 $u(x)$的值在 $0\leqslant x\leqslant a$ 区间中为 $u_1(x)$,在区间$-b\leqslant x<0$ 中为 $u_2(x)$,由式（3-51）得到

$$\frac{\mathrm{d}^2u_1}{\mathrm{d}x^2}+2\mathrm{i}k\frac{\mathrm{d}u_1}{\mathrm{d}x}-\left(k^2-\alpha^2\right)u_1=0\ (0\leqslant x<a) \qquad (3\text{-}53)$$

$$\frac{\mathrm{d}^2u_2}{\mathrm{d}x^2}+2\mathrm{i}k\frac{\mathrm{d}u_2}{\mathrm{d}x}-\left(k^2-\beta^2\right)u_2=0\ (-b\leqslant x<0) \qquad (3\text{-}54)$$

式中

$$\beta=\left[\frac{2m(\varepsilon-V_0)}{\hbar^2}\right]^{1/2} \qquad (3\text{-}55)$$

方程式（3-53）和方程式（3-54）的解分别为

$$u_1(x)=A\exp[\mathrm{i}(\alpha-k)x]+B\exp[-\mathrm{i}(\alpha+k)]\ (0\leqslant x<a) \qquad (3\text{-}56)$$

$$u_2(x)=B\exp[\mathrm{i}(\beta-k)x]+D\exp[-\mathrm{i}(\beta+k)]\ (-b\leqslant x<0) \qquad (3\text{-}57)$$

式中，A、B、C、D 为积分常数。

由量子力学可知，波函数一次微商在 $x=0$ 和 $x=a$ 处必须连续。另外，由于周期性，$u(x)$和 $\mathrm{d}u/\mathrm{d}x$ 在 $x=a$ 和 $x=-b$ 处具有同样的值。因此，边界条件为 $u_1(0)=u_2(0)$，$u_1'(0)=u_2'(0)$，$u_1(a)=u_2(-b)$，$u_1'(a)=u_2'(-b)$。将式（3-56）和式（3-57）代入这些边界条件，且使常数 A、B、C、D 具有非零解，最终求得 ε 和的 k 关系为

$$-\frac{\alpha^2+\beta^2}{2\alpha\beta}\sin(\alpha a)\sin(\beta b)+\cos(\alpha a)\cos(\beta b)=\cos[k(a+b)] \qquad (3\text{-}58)$$

当 $0<\varepsilon<V_0$ 时，根据式（3-55）β 是个纯虚数，令 $\beta=\mathrm{i}\gamma$，由数学公式 $\cos(\mathrm{i}x)=\cosh x$ 和 $\sin(\mathrm{i}x)=\mathrm{i}\sinh x$ 式（3-58）变为

$$\frac{\gamma^2-\alpha^2}{2\alpha\gamma}\sinh(\gamma b)\sin(\alpha a)+\cos(\gamma b)\cos(\alpha a)=\cos[k(a+b)] \qquad (3\text{-}59)$$

为进一步简化，设 V_0b 为常数，使 V_0 趋于无穷大，那么，b 趋于零。由于 $V_0\gg\varepsilon$，则 $\sinh(\gamma b)\approx\gamma b$，$\cosh(\gamma b)\approx 1$，式（3-59）变为

$$P\frac{\sin(\alpha a)}{\alpha a}+\cos(\alpha a)=\cos(ka) \tag{3-60}$$

式中，P 表示 $V_0 b$ 的大小，$P = mV_0ba/\hbar^2$。

当波动方程式（3-48）具有式（3-49）形式的解时，其能量本征值 ε 可由解式（3-60）求得。把式（3-60）的左边作为 αa 的函数用图形表示就得到图 3-12 的曲线。根据定义，波数是实数，而且$-1\leqslant\cos(ka)\leqslant 1$，所以式（3-60）左边的值在+1 和−1 之间，只有图 3-12 中打斜线区域所含的 αa 才存在使左边和右边相等的解。α 代表了能量 ε，根据式（3-60），特定的 ε 才能作为方程式（3-48）许可的本征值。由图 3-12 可以看出，许可的能量状态只存在于打斜线的带状区域里。于是，把存在许可的能量状态和不存在许可的能量状态的带状能量区域称为能带。含有许可的能量状态的能带称为许可带，不存在许可的能量状态的区域称为禁带。在 $0\leqslant ka\leqslant\pi$ 的范围内改变 k，则式（3-60）的右边 $\cos(ka)$在+1 和−1 之间变化，αa 在图 3-12 的 11′区间内变化。同样，k 在 $\pi\leqslant ka\leqslant 2\pi$ 之间变化时，αa 在 22′区间内变化。这样便可求出图 3-13 的粗实线所示的 $\alpha-k$ 的关系即 $\varepsilon-k$ 的关系。由图 3-13 可以看出，在 $\cos(ka)=\pm1$ 即 $k=\pm(\pi/a)n$(n 为正整数)处出现禁带。随着能量的增加，许可带的宽度增大，而禁带的宽度变窄，逐渐趋近自由电子的情形。在式（3-60）中，当参数 P 趋近于零时，$\cos(\alpha a)\approx\cos(ka)$，所以利用式（3-52）可以看出，$\varepsilon-k$ 的

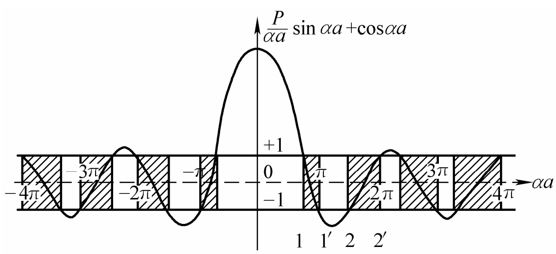


图 3-12　$P/(\alpha a)\sin(\alpha a)+\cos(\alpha a)$对 αa 的函数图（斜线区域存在允许能量状态）

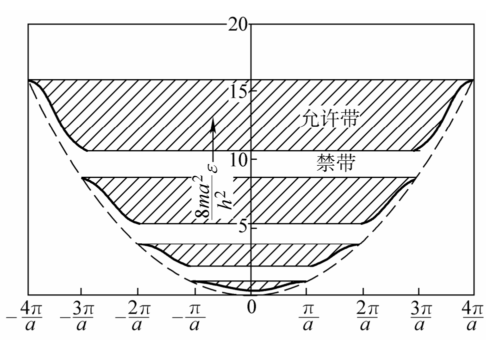


图 3-13　克勒尼希—彭尼模型的能量与波数的关系（$P=3\pi/2$）（虚线表示自由电子模型的结果）

关系趋近于自由电子的情况。另一方面，当 P 为无穷大时，由图 3-12 可以看出，许可带的宽度趋于零，电子的能量状态和孤立原子一致。

3.4 能带理论

由克勒尼希—彭尼模型已经知道晶体中存在着能带。但是，对于晶格的周期势场来说，这个模型过于简单，不适于用来精确描述实际晶体中电子的状态，即使对于周期势场也无法精确求解，所以实际上需要根据情况作适当假定来近似求解。其中一种近似方法称为“准自由电子近似”，它假定电子的动能远大于晶格的周期性势场；另一种近似方法称为“紧束缚电子近似”，它把电子波函数近似为原子轨道函数的线性组合。

3.4.1 准自由电子近似

1. 一维晶体准自由电子近似

假定晶格周期势场对电子的作用很弱，电子几乎是自由地运动着，这时可以认为电子的行为和量子自由电子论中讨论的自由电子相同。但是，又有一点特殊，当电子波的波长 λ（或波数 k）满足布喇格衍射条件时电子将受到由晶格势场引起的布喇格反射而形成驻波，不能在晶体中传播。例如，晶格常数为 a 的一维晶体（图 3-14），当波数为 $k=\pi n/a$（n 为整数）的电子波 $\exp(ikx)$入射时，由相邻原子引起的两个反射波的光程差为 $2a$，相位差为 $k2a=2n\pi$，两个反射波同相位，满足布喇格反射条件式而互相增强。所以，这个波数的电子受到全反射而不能在晶体中行进，具有与此波数相应能量的电子在晶体中不存在。

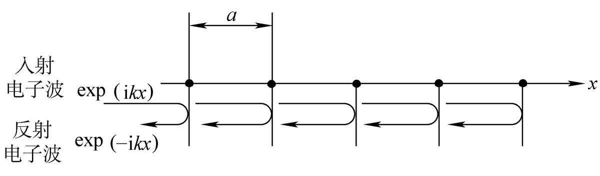


图 3-14 一维晶格中电子波 $\exp(ikx)$的布喇格反射

下面更详细地讨论一下这一点。一维晶体（晶格常数为 a，长有 N 个晶格常数）中，电子的波动方程为

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2\psi_k(x)}{\mathrm{d}x^2}+V(x)\psi_k(x)=\varepsilon_k\psi_k(x) \tag{3-61}$$

脚标 k 表示波函数 $\psi(x)$是由分立的波数 k 确定的（波函数 $\psi_k(x)$布格赫函数，波数 $k = 2\pi g/(Na)$，g 是 $0\leqslant g\leqslant N-1$ 区间中的整数。）。$V(x)$为晶格的周期势场，在准自由电子近似中认为 $\varepsilon_k\gg V(x)$。因为 $V(x)$具有和晶格相同的周期 a，所以可以展开

为如下的傅里叶级数

$$V(x)=\sum_{n=-\infty}^{\infty}V_n\exp\left(-\frac{\mathrm{i}2\pi nx}{a}\right)=\gamma\sum_{n=-\infty}^{\infty}v_n\exp\left(-\frac{\mathrm{i}2\pi nx}{a}\right) \tag{3-62}$$

式中展开系数 $V_n=\gamma v_n$ 为

$$V_n=\gamma v_n=\frac{1}{2}\int_0^2 V(x)\exp\left(\frac{\mathrm{i}2\pi nx}{a}\right)\mathrm{d}x \tag{3-63}$$

若把能量原点选在 $V(x)$的平均值$(1/a)\int_0^a V(x)\mathrm{d}x$ 为零处，则当 $n=0$ 时 $V_0=0$。另外，系数 γ 为远小于 1 的常数，对于自由电子 $\gamma=0$。

满足式（3-61）的波函数是式（3-45）的布格赫函数，所以具有和晶格势场相同周期的函数 $u_k(x)$也能展开为如下的傅里叶级数

$$u_k(x)=\sum_{n=-\infty}^{\infty}b_n\exp\left(-\frac{\mathrm{i}2\pi nx}{a}\right) \tag{3-64}$$

因此

$$\psi_k(x)=u_k(x)\exp(\mathrm{i}kx)=\exp(\mathrm{i}kx)\sum_n b_n\exp\left(-\frac{\mathrm{i}2\pi nx}{a}\right) \tag{3-65}$$

对于自由电子，$V(x)=0$（在式（3-62）中 $\gamma=0$）波函数 $\psi_k(x)=b_0\exp(\mathrm{i}kx)$，因此式（3-65）的布格赫函数近似地可写成

$$\psi_k(x)=b_0\exp(\mathrm{i}kx)+\gamma\left[\exp(\mathrm{i}kx)\sum_{n\neq0}b_n\exp\left(-\frac{\mathrm{i}2\pi nx}{a}\right)\right] \tag{3-66}$$

式（3-66）满足布格赫条件式（3-46）。

把式（3-62）、式（3-66）代入式（3-61），最终求得二级近似的能量本征值 ε_k 为

$$\begin{aligned}\varepsilon_k&=\frac{\hbar^2k^2}{2m}+\gamma^2\sum_{n\neq0}\frac{b_nv_n^*}{b_0}=\frac{\hbar^2k^2}{2m}+\gamma^2\sum_{n\neq0}\frac{v_nv_n^*}{\hbar^2k^2/(2m)-\hbar^2k_n^2/(2m)}\\&=\frac{\hbar^2k^2}{2m}+\gamma^2\sum_{n\neq0}\frac{|v_n|^2}{\hbar^2k^2/(2m)-[\hbar^2/(2m)](k-2\pi n/a)^2}\end{aligned} \tag{3-67}$$

式中 v_n^*为 v_n 的共轭复数，$k_n=k-2n\pi/a$。当 k^2 不趋近任一个 k_n^2 时，式（3-67）能够给出准确的近似解。但是，当 $k^2\approx k_n^2$ 时，尽管 γ 非常小，式（3-67）的右边第二项仍是发散的。这时不能把式（3-66）作为波函数的近似。当 $k^2\approx k_n^2$ 时，也即 $|k|=|k-2\pi n/a|$或者 $k=n\pi/a$ 时，可最终求得能量本征值 ε_k 为

$$\varepsilon_k=\frac{\hbar^2}{2m}\left(\frac{\pi n}{a}\right)^2\pm|V_n| \tag{3-68}$$

即满足布喇格衍射条件能量发生 $2\left|V_n\right|$ 的跃变，出现了禁带。由上面的自由电子近似，可以得到如图 3-15 所示的电子能量 ε_k 与波数 k 的关系。

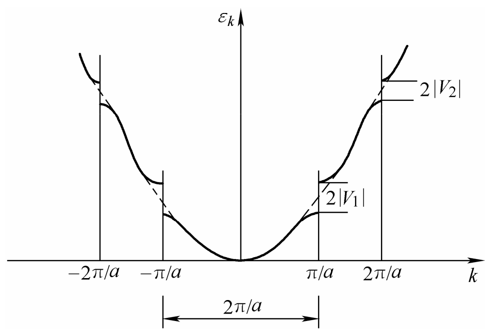


图 3-15 自由电子近似的一维晶格电子能量 ε_k 和波数 k 的关系

2. 晶体中准自由电子近似

（1）$\boldsymbol{k}$ 空间与布里渊（Brillouin）区

上面讨论了一维晶体的 $\varepsilon-k$ 关系及允许带和禁带，现在讨论三维晶体。在三维晶体中，可以建立一个波矢 $\boldsymbol{k}$ 空间来表示电子的能量状态。$\boldsymbol{k}$ 大小等于 $2\pi/\lambda$（波数），方向表示电子波受布喇格衍射前的运动方向。波矢 $\boldsymbol{k}$ 空间由相互垂直的三个波矢分量 k_x、k_y、k_z 的坐标轴组成，与在实际空间中平行于 x、y、z 轴的波运动相对应。一般选择 x、y、z 轴使之与晶体的相应轴平行，例如在立方系中 k_x、k_y、k_z 与单位晶胞相应的三条相互垂直的棱的方向平行。在 $\boldsymbol{k}$ 空间中，等能面上各点的垂直方向即为电子运动的方向。周期场中运动的电子，由于周期排列的金属离子的作用，这时在 $\boldsymbol{k}$ 空间中的等能面有的不是球面。电子运动的方向仍与等能面垂直，但往往与过原点连线方向不一致。

应用 $\boldsymbol{k}$ 空间的概念可以描述在二维及三维晶体中波矢与电子能量的关系。为了方便起见先从二维点阵入手，以后再推广到三维中去。设电子在二维正方点阵中运动，如图 3-16a 代表离子的一个简单正方点阵。图 3-16b 代表相应二维波矢空间，k_x 和 k_y 平行于正方形的两边。在图中画出一条任意直线 $0A$，它代表在相同波运动方向上的一系列电子状态。当 k 值增加时，在 ε_k-k 曲线上响应点出现能量间隙。现假设在图 3-16c 中，线 p_1、p_2、p_3 代表晶体中的离子排列，FG 方向同图 3-16b 中 $0A$ 方向。一般情况下，由 $0A$ 上的波矢来表示其状态的电子将通过这些离子列而很少散射。少数电子的波长能满足布喇格方程而发生强烈反射，电子不能通过点阵。如果要过渡到更高的波矢，能量将有一个陡增，在波矢图 3-16b 中，将有 a_1、a_2、a_3 等一系列波数的点能满足布拉格方程。从 0 到 a_1、a_1 到 a_2、a_2 到 a_3 等波矢区间中能量是连续增加的，但在 a_1、a_2、a_3 等处，能量是急剧增加的。

可以具体计算符合反射条件的 k 值。当波长为 λ 的电子波在晶体中沿某一方

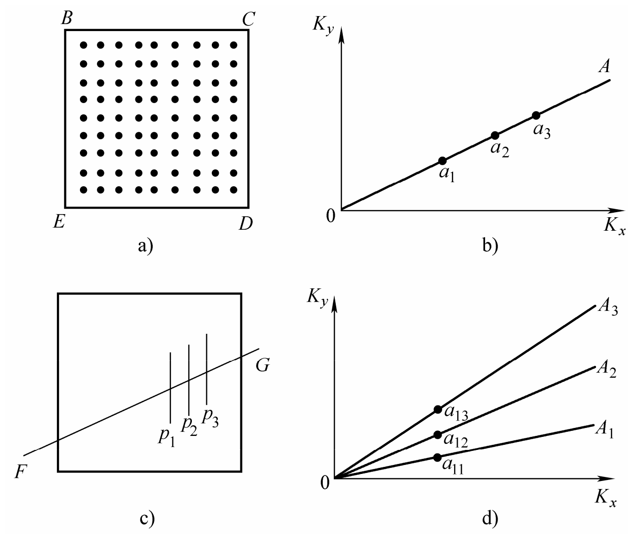


图 3-16　受到反射的 k 值

向运动时，当某一晶面的间距 d 和该晶面与电子波运动方向的夹角 θ 满足布喇格方程 $n\lambda = 2d\sin\theta$ 时，发生全反射。这时，当 $n = 1$ 时，$k = 2\pi/\lambda = \pi/(d\sin\theta)$。对于图 3-16a 所示二维晶格中的 BE、CD 面来说，$d = a$，电子沿 ED 方向运动，$\theta = 90°$，则 $k_x = \pi/a$，即当 $k_x = \pi/a$ 时符合衍射条件；对 BC、ED 面来说，$d = a$，电子沿 EB 方向运动，$\theta = 90°$，当 $k_y = \pi/a$ 时符合衍射条件；对于 BE、CD 面，如果电子运动方向与该组晶面成 45°，即 $\theta = 45°$，则 $k_{45°} = 2^{1/2}\pi/a$。同理可求得各个不同方向符合反射条件的 k 值，如图 3-16d 中的 a_{11}、a_{12}、a_{13} 所示。连接这些点形成的多边形（包括多边形的内部）称为二维结构的第一布里渊（Brillouin）区。同样，图 3-16b 中的 a_2 是第二个多边形边界上的一点，第二个多边形为第二布里渊区边界。布里渊区的意义是：它的边界上 k 值的能量是不连续变化的，而布里渊区内部能量随 k 值连续变化。每个晶体结构有自己特定的布里渊区，它由晶体结构的几何性质来决定的。一个简单正方点阵的第一、第二、第三布里渊区如图 3-17 所示。

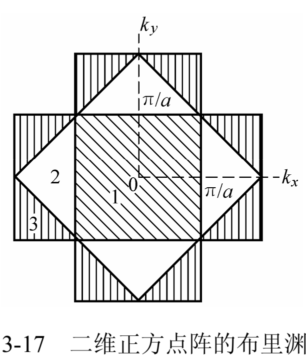


图 3-17　二维正方点阵的布里渊区

同样的原理也适用于三维晶体，只要把二维波矢代之以三维波矢。在三维波矢空间中，找出任意方向上的能态符合布拉格反射的波数临界值点。连接这些点，形成三维的布里渊区。简单立方点阵晶体，其第一布里渊区是一个以 $2\pi/a$ 为边的立方体，重心与原点重合。面心立方点阵与体心立方点阵的第一布里渊区

分别如图 3-18 和图 3-19 所示。

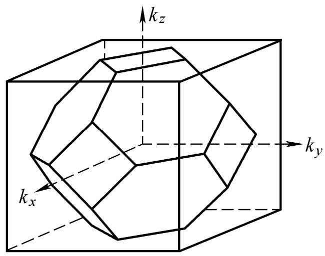


图 3-18　面心立方点阵的第一布里渊区（图中的多面体）

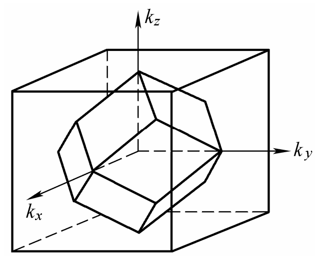


图 3-19　体心立方点阵的第一布里渊区（图中的多面体）

应该说明的是，布里渊区不是真实的空间晶体，它是 k 空间的图像。

（2）布里渊区与能带

在波矢 $\boldsymbol{k}$ 空间中，电子能量相等的点(k_x, k_y, k_z)连成的面称为等能面。可根据电子的 $\varepsilon-k$ 关系画出等能面。现在考查一下电子在 $\boldsymbol{k}$ 空间中的布里渊区里的分布。设想在离子构成的金属晶体中逐点加入“准自由电子”，电子将自低能态向高能态逐步填充。在二维正方点阵的情况下，准自由电子近似的等能线如图 3-21 所示。K 值较小能量低的等能线 1、2 距布里渊区边界较远，是以原点为圆心的圆。这是因为波数与临界值（满足布喇格衍射条件的值）相差较大的电子与自由电子近似。这些电子没有周期场给予它们的方向性，所以不论运动方向如何，都有同样的 $\varepsilon-k$ 关系。当 k 值继续增大（等能线 3），等能线开始偏离圆形，它在接近布里渊区边界处向外突出。这是因为在边界处周期场的影响显著。另外从 $\varepsilon-k$ 曲线上也可看出，在接近临界 k 值处，$\varepsilon-k$ 曲线偏离到自由电子的抛物线以下，ε 随 k 值缓慢地增加。因此在 $\boldsymbol{k}$ 空间的这个区域里，能量从一个等能线增加到另一个等能线，k 值有一个较大的增值。再高的等能线如等能线 4、5，已与布里渊区边界相交，表明具有这个范围的能量的电子状态位于布里渊区的角部，处于布里渊区角顶的能级在这个布里渊区中能量最高。因为在布里渊区的边界上能量是不连续的，等能面不能穿过布里渊区边界。若把第二个布里渊区等能量的几条线也画出，就是如图 3-20 所示的第一布里渊区以外的那几条。布里渊区边界虽有能量垒存在，但整个晶体不一定有禁带。若第一区 k_y 方向最高能级 P 为 3.0eV，这个方向的垒高为 2.0eV，则第二区最低能级 R 为 5.0eV。若第一布里渊区的最高能级 Q 为 4.5eV，这种情况下，第一布里渊区的最高能量低于第二布里渊区的最低能量，则整个晶体就有禁带存在，第一、第二布里渊区能带是分立的。若第一布里渊区的最高能级 Q 为 5.5eV，在这种情况下，整个晶体便没有禁带，第一和第二区能带重叠。两种情况电子在布里渊区的填充如图 3-21 所示。图中 1、2、3、4 表示布里渊区

的各个充填阶段，已充填的区域画有影线。图 3-21a 为具有禁带的情况；图 3-21b 为无禁带的情况。

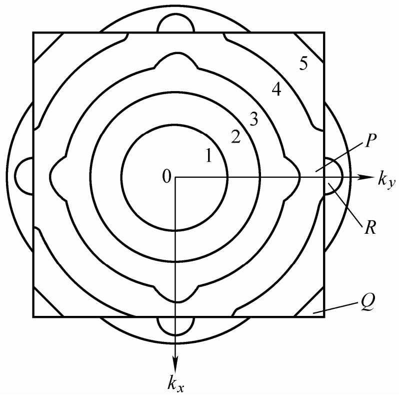


图 3-20　二维正方点阵第一、二布里渊等能线

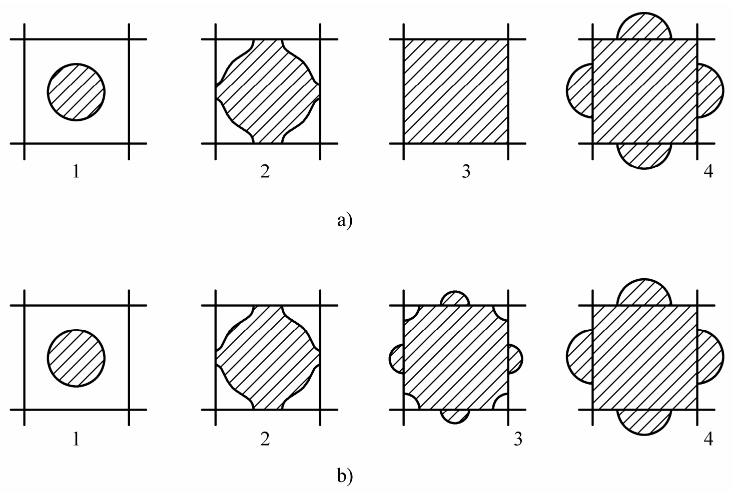


图 3-21　电子填充布里渊区的不同阶段

a）无重叠布里渊区　b）有重叠布里渊区

上述情况如用 $\varepsilon-k$ 曲线表示，则如图 3-22 所示。图 3-22 是把几个不同方向的电子波 $\varepsilon-k$ 曲线叠放在一个平面图上而得出的。在图 3-22a 中，可看到第一布里渊区中最高点 c 低于第二布里渊区能量最低点 a'，这说明不存在能带重叠。图 3-22b 示出的是能带重叠的情况，即虽在任意一个方向上无能带重叠，但各个方向组合在一起，第一布里渊区的能量最高点 c 高于第二布里渊区的能量最低点 a'，故该晶体属于能带重叠型的。是否存在能带重叠对晶体性能影响极大。

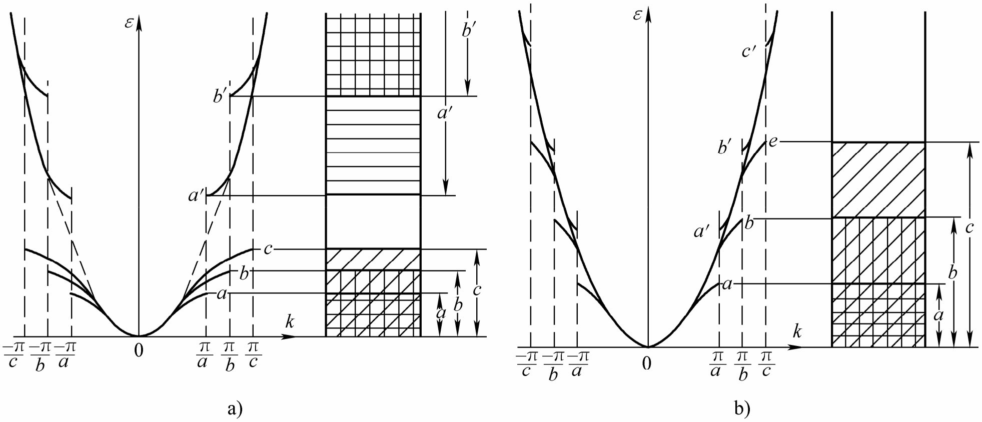


图 3-22　不同方向的 $\varepsilon-k$ 曲线叠放在一个平面图上的表现

a）无能带重叠（有禁带）　b）有能带重叠（无禁带）

（3）电子状态密度分布

当电子充填 $\boldsymbol{k}$ 空间的最初阶段，状态密度 $\rho(\varepsilon)$与能量 ε 的关系曲线如量子自由电子模型那样呈抛物线。当在 $\boldsymbol{k}$ 空间中表示被占状态的球扩大且接近布里渊区边界的最近点时，$\rho(\varepsilon)-\varepsilon$ 曲线就开始偏离抛物线（图 3-23 中 AB 段）。即当接近布里渊区边界时，ε 随 k 增加较慢，因此状态密度增高较多。在这个范围内，$\rho(\varepsilon)-\varepsilon$ 曲线位于抛物线之上。当这个变形的球继续增大，并终于在最近点碰到布里渊区边界时，$\rho(\varepsilon)$曲线达到最高点 B（对应图 3-22 中 a 点对应能量，即图 3-23 中 a 点能量）。在这以后，只有布里渊区的角落高能量部分可以充填电子，由于角落部分的状态数少，因此状态密度降低。对无能带重叠（有禁带）的情况，当布里渊区完全填满时，曲线到达 c，$\rho(\varepsilon)$降为 0。c 代表第一带最后一条曲线的顶部（图 3-22a）。当第一布里渊区填满后，再充填下一个电子状态 a'（图 3-22 中 a' 点对应能量，即图 3-23 中 a'点能量），能量将有个陡增。这时出现了图 3-23 禁带 ca'。此时，再增加电子将充填第二布里渊区（图 3-23 中 $a'C$），这时要求具有更高的能量。

第二种是布里渊区产生重迭的情况（参见图 3-22b）。其能态密度曲线如图 3-24。进入第一布里渊区最初的电子和上述情况一样。当充填状态的费米面到达第一布里渊区的边界时，在 A 点发生陡降。当到达图 3-24 中的 a'点对应的能量时，电子开始进入第二布里渊区。从图 3-22b 中可看到 a'相当于第二布里渊区的最低能量状态。因为第一布里渊区 c 状态的能量高于第二布里渊区的 a'状态的能量，当继续充填更多电子时，它们充填第一布里渊区的剩余状态和第二布里渊区的低能状态，在 $\rho(\varepsilon)-\varepsilon$ 曲线上 $\rho(\varepsilon)$表示的是两个重叠区纵坐标的和，如图 3-24 中

BCD 所示。所以 $\rho(\varepsilon)-\varepsilon$ 曲线在 a'到 c 之间突然升高。到达 c 点对应的能量时，第一布里渊区完全填满。以后的电子继续充填第二布里渊区。如果第二布里渊区和第三布里渊区重叠，也产生同样的效应，在 $\rho(\varepsilon)-\varepsilon$ 曲线上有一系列的峰谷，其形状取决于晶体结构和区域之间能隙大小。

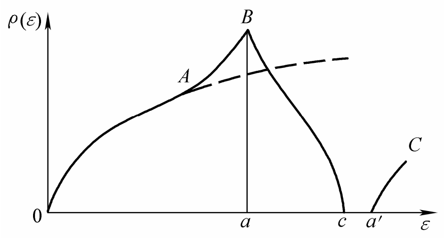


图 3-23　无能带重叠（有禁带）$\rho(\varepsilon)-\varepsilon$ 曲线

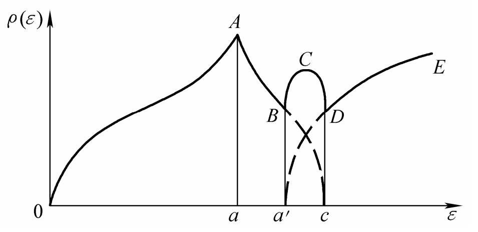


图 3-24　有能带重叠（无禁带）$\rho(\varepsilon)-\varepsilon$ 曲线

用长波（波长约在 100Å（10nm））X 线谱法（软 X 线谱）可以测定金属及合金在费米能以下的 $\rho(\varepsilon)-\varepsilon$ 曲线，这是一种研究金属及合金电子结构的方法。

（4）费密面

金属中的电子自低能级排起，直到全部价电子占据了相应的能级为止。在 0K 时低于费密能 $\varepsilon_F(0)$的能级全被电子填满，高于费密能 $\varepsilon_F(0)$的能级全空着。如果与 $\varepsilon_F(0)$相对应的波数为 k_F，对于自由电子来说，以 k_F 为半径的球体内填满电子，球外均为空能级，这个球的表面称为费密面。对于在晶体周期势场作用下的准自由电子来说，其费密面有些不是球面，而呈各种复杂的多面体的外形轮廓。对于简单立方金属，当电子少时，电子排布在第一布里渊区的原点附近，此时费密面为球面；但当电子逐步增多，电子在 k 空间的排布已接近布里渊区边界时，此时 $\varepsilon-k$ 面已偏离球面而向外凸出，成为一个多面体。在与布里渊区边界相交时，费密面的变形更为显著。

对于具有体心立方晶体结构的碱金属，如 Li、Na 和 K 等，它们的价带是半满的，费密面距布里渊区边界还较远，因此这些金属的费密面是个球面。而具有面心立方晶体结构的贵金属，如 Cu、Ag、Au 等，其价带虽为半满，但沿[111]方向的费密面离布里渊区边界较近，因而发生强烈变形。

费密面这个概念的建立，其意义在于对固体的若干物理现象的理解。如导电，导热、热容，霍尔效应等均与费密面的概念有密切关系。即实际参加物理过程的电子并不是所有价电子，而只是靠近费密面附近的那些电子才是物理过程参加者，因此研究这些物理现象时，就要重点考查费密面附近电子的行为。

3.4.2　紧束缚电子近似（原子轨道线性组合法）

晶体中原子间距 a 较大，晶格势变化显著，在原子附近电子受自身原子的束

缚较紧，不容易产生共有化运动。近原子区，电子的行为同孤立原子中的电子行为相似，晶体波函数也相应地接近于孤立原子的波函数，因此可用原子轨道函数的线性组合构成紧束缚近似的电子波函数 $\psi(\boldsymbol{k}, \boldsymbol{r})$

$$\psi_i(\boldsymbol{k},\boldsymbol{r}) = \sum_n C_n(\boldsymbol{k})\varphi_i(\boldsymbol{r}-\boldsymbol{R}_n) \tag{3-69}$$

作为晶体波函数，它必须具有布洛赫函数的形式，并且构成正交归一系，满足这样条件的紧束缚近似晶体波函数 $\psi_i(\boldsymbol{k}, \boldsymbol{r})$可写成

$$\psi_i(\boldsymbol{k},\boldsymbol{r}) = \sum_n \frac{1}{\sqrt{N}} \mathrm{e}^{\mathrm{i}\boldsymbol{k}\boldsymbol{R}_n} \varphi_i(\boldsymbol{r}-\boldsymbol{R}_n) \tag{3-70}$$

上式中 $\varphi_i(\boldsymbol{r}-\boldsymbol{R}_n)$代表位于格点 $\boldsymbol{R}_n$ 的孤立原子波函数，下标代表原子轨道。$\psi_i(\boldsymbol{k}, \boldsymbol{r})$中的下标与 $\varphi_i(\boldsymbol{r}-\boldsymbol{R}_n)$的下标对应。$N$ 为晶体总原子数，系数 $1/N^{1/2}$ 为轨道函数重迭可以忽略时的归一化系数。现在讨论非简并的 s 态，这时相应的原子波函数只有一个 $\varphi_i(\boldsymbol{r}-\boldsymbol{R}_n)$，它满足

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + V(\boldsymbol{r}-\boldsymbol{R}_n)\right]\varphi_i(\boldsymbol{r}-\boldsymbol{R}_n) = \varepsilon_i\varphi_i(\boldsymbol{r}-\boldsymbol{R}_n) \tag{3-71}$$

E_i 代表 s 态的原子能级。这个模型称为紧束缚近似。

如图 3-25 所示，晶格中在 $\boldsymbol{R}_n$ 附近任意点 A 的电子势能为

$$V(\boldsymbol{r}) = V(\boldsymbol{r}) + V(\boldsymbol{r}-\boldsymbol{R}_n) - V(\boldsymbol{r}-\boldsymbol{R}_n) = V(\boldsymbol{r}-\boldsymbol{R}_n) + V'(\boldsymbol{r}-\boldsymbol{R}_n) \tag{3-72}$$

$V(\boldsymbol{r}-\boldsymbol{R}_n)$为位于 $\boldsymbol{R}_n$ 的原子在 $\boldsymbol{r}$ 点产生的势能。$V'(\boldsymbol{r}-\boldsymbol{R}_n) = V(\boldsymbol{r})-V(\boldsymbol{r}-\boldsymbol{R}_n)$，它代表扣除了位于 $\boldsymbol{R}_n$ 以外的所有其他原子在 $\boldsymbol{r}$ 点产生的电子势能迭加，即

$$V'(\boldsymbol{r}-\boldsymbol{R}_n) = \sum_j V(\boldsymbol{r}-\boldsymbol{R}_j) \tag{3-73}$$

式中的求和项不包括 $\boldsymbol{R}_j = \boldsymbol{R}_n$。

紧束缚电子满足的薛定谔方程为

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + V(\boldsymbol{r})\right]\psi_i(\boldsymbol{k},\boldsymbol{r}) = H\psi_i(\boldsymbol{k},\boldsymbol{r}) = \varepsilon_i(\boldsymbol{k})\psi_i(\boldsymbol{k},\boldsymbol{r}) \tag{3-74}$$

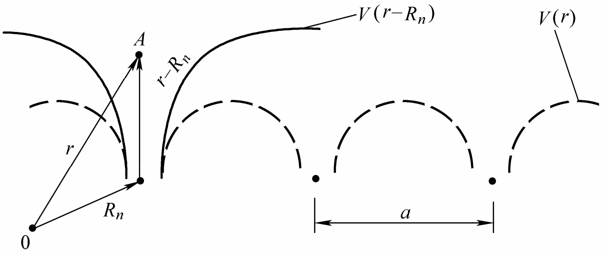


图 3-25　紧束缚近似中的晶格势场

这里 H 为哈密顿算符。式（3-74）即

$$\left[-\frac{\hbar^2}{2m}\nabla^2+V(\boldsymbol{r}-\boldsymbol{R}_n)+V'(\boldsymbol{r}-\boldsymbol{R}_n)\right]\psi_i(\boldsymbol{k},\boldsymbol{r})=\varepsilon_i(\boldsymbol{k})\psi_i(\boldsymbol{k},\boldsymbol{r}) \tag{3-75}$$

将式（3-70）、式（3-72）代入式（3-74），再乘 $\psi_i(\boldsymbol{k}, \boldsymbol{r})$的共轭复数 $\psi_i^*(\boldsymbol{k}, \boldsymbol{r})$，对整个体积 τ 积分，求得能量 $\varepsilon_i(\boldsymbol{k})$为

$$\varepsilon_i(\boldsymbol{k})=\frac{\int\psi_i^*H\psi_i\mathrm{d}\tau}{\int\psi_i^*\psi_i\mathrm{d}\tau} \tag{3-76}$$

$$\int\psi_i^*\psi_i\mathrm{d}\tau=\frac{1}{N}\sum_n\sum_m\mathrm{e}^{\mathrm{i}\boldsymbol{k}(\boldsymbol{R}_n-\boldsymbol{R}_m)}\int\varphi_i^*(\boldsymbol{r}-\boldsymbol{R}_m)\varphi_i(\boldsymbol{r}-\boldsymbol{R}_n)\mathrm{d}\tau \tag{3-77}$$

当 $\boldsymbol{R}_n=\boldsymbol{R}_m$ 时，上式的重迭积分值最大，如略去 $\boldsymbol{R}_n\neq\boldsymbol{R}_m$ 时的重迭积分值，则 $\int\psi_i^*\psi_i\mathrm{d}\tau=1$。因此

$$\varepsilon_i(\boldsymbol{k})=\int\psi_i^*H\psi_i\mathrm{d}\tau=\frac{1}{N}\sum_n\sum_m\mathrm{e}^{\mathrm{i}\boldsymbol{k}(\boldsymbol{R}_n-\boldsymbol{R}_m)}\int\varphi_i^*(\boldsymbol{r}-\boldsymbol{R}_m)H\psi_i(\boldsymbol{r}-\boldsymbol{R}_n)\mathrm{d}\tau \tag{3-78}$$

或

$$\begin{aligned}\varepsilon_i(\boldsymbol{k})=&\frac{1}{N}\sum_n\sum_m\mathrm{e}^{\mathrm{i}\boldsymbol{k}(\boldsymbol{R}_n-\boldsymbol{R}_m)}\int\varphi_i^*(\boldsymbol{r}-\boldsymbol{R}_m)\left[-\frac{\hbar^2}{2m}\nabla^2+V(\boldsymbol{r}-\boldsymbol{R}_n)\right]\varphi_i(\boldsymbol{r}-\boldsymbol{R}_n)\mathrm{d}\tau\\&+\frac{1}{N}\sum_m\int\varphi_i^*(\boldsymbol{r}-\boldsymbol{R}_m)V'(\boldsymbol{r}-\boldsymbol{R}_m)\varphi_i(\boldsymbol{r}-\boldsymbol{R}_m)\mathrm{d}\tau\\&+\frac{1}{N}\sum_m^{n\text{近邻}}\sum_n\mathrm{e}^{\mathrm{i}\boldsymbol{k}(\boldsymbol{R}_n-\boldsymbol{R}_m)}\int\varphi_i^*(\boldsymbol{r}-\boldsymbol{R}_m)V'(\boldsymbol{r}-\boldsymbol{R}_n)\varphi_i(\boldsymbol{r}-\boldsymbol{R}_n)\mathrm{d}\tau\\=&\varepsilon_i-\alpha-\beta\sum_{R_m}^{\text{近邻}}\mathrm{e}^{-\mathrm{i}\boldsymbol{k}(\boldsymbol{R}_m-\boldsymbol{R}_n)}\end{aligned} \tag{3-79}$$

式中 ε_i 为原子能级，α 和 β 的积分表达式为

$$-\alpha=\int\varphi_i^*(\boldsymbol{r}-\boldsymbol{R}_m)V'(\boldsymbol{r}-\boldsymbol{R}_m)\varphi_i(\boldsymbol{r}-\boldsymbol{R}_m)\mathrm{d}\tau \tag{3-80}$$

$$-\beta=\int\varphi_i^*(\boldsymbol{r}-\boldsymbol{R}_m)V'(\boldsymbol{r}-\boldsymbol{R}_n)\varphi_i(\boldsymbol{r}-\boldsymbol{R}_n)\mathrm{d}\tau \tag{3-81}$$

$\varepsilon_i(\boldsymbol{k})$表式中，由于第三项的出现（$\beta$ 表示重迭积分），原子的分立能级展宽成能带。图 3-26 代表紧束缚近似法获得的能带示意图。图中表示当原子彼此远离时（重迭积分小），能带变窄，极限情况下，它趋近于孤立原子的能级。当构成晶体的 N 个原子靠得很近时（重迭积分较大），孤立原子能级展宽并分裂成 N 个接近连续的密集能级，形成能带。

因为式（3-80）和式（3-81）的积分计算极为麻烦，有时把 α 和 β 作为常数来

处理。

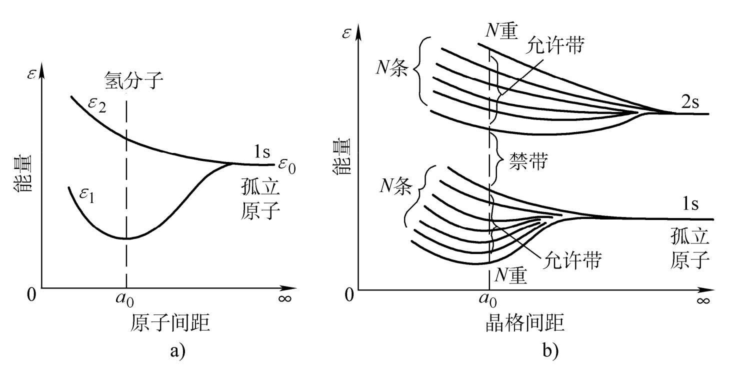


图 3-26　能带形成

a）氢分子能级　b）固体能带的形成

下面就边长为 a 的简立方晶格具体地讨论一下式（3-79），因为 $\boldsymbol{R}_m-\boldsymbol{R}_n$ 只取最邻近的原子就行，所以 $\boldsymbol{R}_m-\boldsymbol{R}_n$ 的分量为 $\boldsymbol{R}_m-\boldsymbol{R}_n=(\pm a, 0, 0), (0, \pm a, 0), (0, 0, \pm a)$。如果讨论 s 电子轨道的轨道函数，对于六个最邻近的原子 β 相等，因此式（3-79）为

$$\varepsilon_i(\boldsymbol{k})=\varepsilon_i-\alpha-2\beta(\cos k_x a+\cos k_y a+\cos k_z a) \qquad (3\text{-}82)$$

若把式（3-82）沿简单立方晶格的[100]及[111]方向用图表示出来，就形成了如图 3-27 所示的能带。$\varepsilon_i(\boldsymbol{k})$在布里渊区中心 $k=(0, 0, 0)$处取最小值 $\varepsilon_i(0, 0, 0)=\varepsilon_i-\alpha-6\beta$；在[111]方向的布里渊区边界取最大值 $\varepsilon_i(1, 1, 1)=\varepsilon_i-\alpha+6\beta$。在 k 很小的范围内，可以把式（3-82）的 cos 项展开为

$$\varepsilon_i(\boldsymbol{k})=\varepsilon_i-\alpha-6\beta+\beta a^2({k_x}^2+{k_y}^2+{k_z}^2)=\varepsilon_i-\alpha-6\beta+\beta a^2k^2 \qquad (3\text{-}83)$$

这时，与自由电子一样，等能面为球面。另外，在[111]方向的布里渊区边界附近，也可以把式（3-82）的 cos 项展开。令 $k_x'=\pi/a-k_x$，当 k_x'很小时，$\cos(\pi-k_x'a)\approx-1+(k_x'a)^2/2$，因此在[111]方向能带顶附近

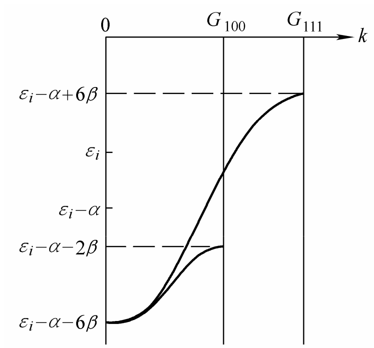


图 3-27　由紧束缚近似求出的简单立方晶格沿[100]和[111]方向的能带

$$\varepsilon_i(\boldsymbol{k}) = \varepsilon_i - \alpha - 6\beta - \beta a^2({k'_x}^2 + {k'_y}^2 + {k'_z}^2) = \varepsilon_i - \alpha - 6\beta - \beta a^2 k'^2 \quad (3\text{-}84)$$

因为波数 k'是在[111]方向布里渊区边界测得的值，所以等能面是以布里渊区的[111]方向的顶点为中心的球面。图 3-28 表示二维正方晶格的 $\boldsymbol{k}$ 空间中的等能线。在准自由电子近似的情况下，如图 3-28b 所示，$\varepsilon_i(\boldsymbol{k})$与 k^2 成正比的范围扩展到波数较大的区域。

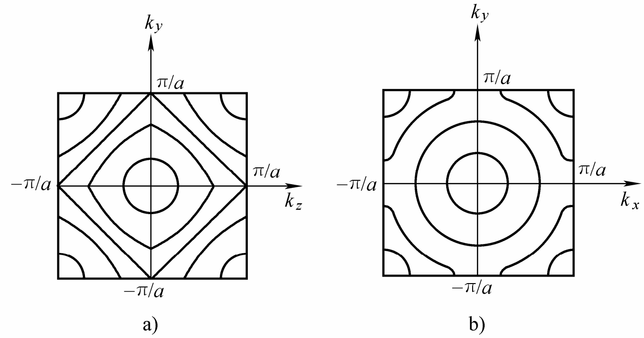


图 3-28　二维晶格第一布里渊区中的等能现

a）紧束缚近似　b）准自由电子近似

锗和硅与金刚石有相同的结构。由 N 个碳原子组成的金刚石晶体，其能带与原子间距的关系如图 3-29 所示。这是因为原子间距变小时，相邻原子轨道的重迭增加，因此式（3-79）的 α、β 增大，能带宽度变宽。当原子间距从孤立状态逐渐减小时，在孤立状态中，包含自旋 2 重简并的 s 状态及 6 重简并的 p 状态，都因为和邻近原子轨道函数重迭使简并解除而分别形成了能带。当原子间距减小到 a' 时，s 状态和 p 状态共有化形成新的轨道（sp^3 混合轨道），出现了两个新的能带，这两个能带分别包含 4N 个能级。因一个原子中包含有 4 个价电子，所以总数 4N 个价电子完全填满了下面的能带。

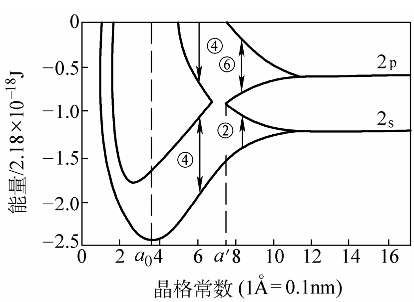


图 3-29　金刚石结构能带（O 内的数子表示单原子时的状态数）

准自由电子近似和紧束缚电子近似两种方法互为补充，对价电子近似自由电

子的，如碱金属和贵金属的价电子，准自由电子近似较为合适。而当电子比较紧密地束缚于它们的原子时，如过渡族金属亚层电子，则紧束缚电子近似更合适。

能带理论继承和发展了电子论，弥补了量子自由电子论的不足，在解释导电性、铁磁性、相结构以及结合力等方面都取得成功。在金属各种物理性能的讨论中，能带理论是个重要的基础。

3.5 能带中的电子运动

在周期势场中运动的电子的波函数，由遍及整个晶体的布洛赫函数给出。但是，在处理输运现象时，把电子作为粒子来描述最方便。设电子占有的空间范围为 Δx，根据测不准原理 $\Delta x \Delta p_x \approx h$，动量也有个变化范围 $\Delta p_x \approx h/\Delta x$。电子的粒子性可以由波数相近的许多束波（布洛赫波）的迭加来表示。波数不同的许多束波迭加可以得到如图 3-30 所示局限在一定空间的波，称为波包。作为粒子的电子其运动速度是波包振幅最大的点（即存在几率最大的点）的速度，而不是构成波包的各个傅里叶成分的波的相速度 ω/k。可以把德布罗意波组成波包用粒子的观点讨论波包的运动。这个方法可用于讨论晶体电子在外场作用下的运动规律。

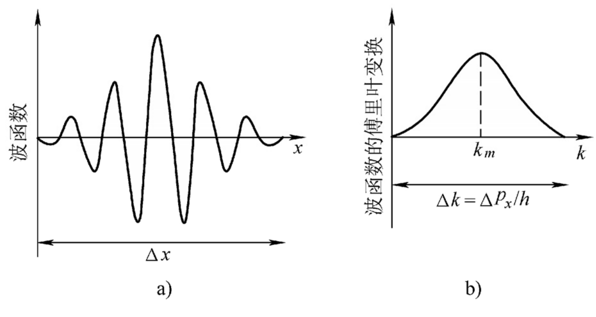


图 3-30 波包

a）波包空间扩展范围 b）波数扩展范围

以 k_0 为中心，将波矢在 $\Delta\boldsymbol{k}$ 范围内变化的布洛赫波组成波包，在 $\Delta\boldsymbol{k}$ 比布里渊区的限度 $2\pi/a$ 小得多（也就是波包中心展宽的范围 $\Delta x \gg a$，a 为晶格间距）的条件下，布洛赫电子的运动可以看作以 k_0 为中心的波包运动。这个前提下，晶体电子可用类经典粒子所具有的速度、准动量和能量等经典量描述，并可用有效质量 m^*来概括与晶格有相互作用的电子对外力 F 的响应，把电子在外力作用下的加速度 $\boldsymbol{a}$ 与外力 $\boldsymbol{F}$ 间的关系写成类似牛顿第二定律的形式，即 $\boldsymbol{F} = m^*\boldsymbol{a}$。

3.5.1 电子速度 $v(\boldsymbol{k})$ 与波矢 $\boldsymbol{k}$ 的关系

在第 2 章中已阐明沿某一方向运动的波包的群速度 $u_g = \mathrm{d}\omega/\mathrm{d}k$（$\omega$ 角频率，k

为波数），电子运动的速度等于布洛赫波组成波包的群速度，即电子运动速度 $v(k) = u_g = \mathrm{d}\omega/\mathrm{d}k$。对于物质波，粒子的能量 $E(k) = h\nu = \hbar\omega$，那么对自由电子运动的运动（总能量中只有动能），$v(k)$为

$$v(k)=\frac{\mathrm{d}\omega}{\mathrm{d}k}=\frac{1}{\hbar}\frac{\mathrm{d}\varepsilon(k)}{\mathrm{d}k}=\frac{\hbar k}{m} \tag{3-85}$$

这里 m 为电子的质量。用矢量表示速度的方向和波矢的方向，则

$$\boldsymbol{v}(\boldsymbol{k})=\frac{1}{\hbar}\nabla_{\boldsymbol{k}}\varepsilon(\boldsymbol{k})=\frac{\hbar\boldsymbol{k}}{m} \tag{3-86}$$

∇ 是矢量微分算符。$\boldsymbol{v}(\boldsymbol{k})$沿球形等能面法线方向（$\boldsymbol{k}$ 方向），并与波矢 $\boldsymbol{k}$ 成正比，如图 3-31a 所示。

处于晶体中的电子（布洛赫电子），其运动速度等于波包中心的群速度，它与 $\boldsymbol{k}$ 空间的能量梯度成正比，即

$$\boldsymbol{v}(\boldsymbol{k})=\frac{1}{\hbar}\nabla_{\boldsymbol{k}}\varepsilon(\boldsymbol{k}) \tag{3-87}$$

速度 $\boldsymbol{v}(\boldsymbol{k})$的方向沿等能面的法向。由于晶格周期场势的作用，其等能面并非总是球形等能面，$\boldsymbol{v}(\boldsymbol{k})$与 $\boldsymbol{k}$ 并不总是同向，大小也不一定等于 $\hbar k/m$（但等于 $\hbar k/m^*$，m^*为电子有效质量，后面叙述）。其 $\boldsymbol{v}(\boldsymbol{k})$～$\boldsymbol{k}$ 关系如图 3-31b 所示。

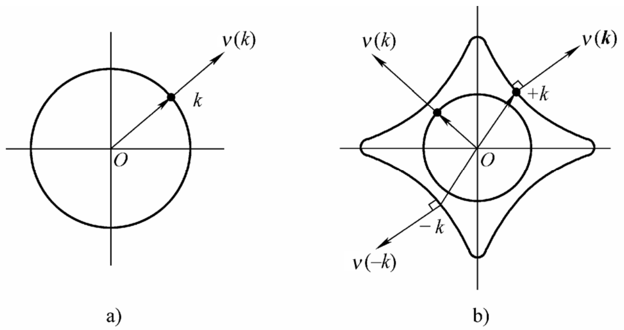


图 3-31　电子的速度矢量

a）自由电子　b）布洛赫电子

电子的能量 $\varepsilon(\boldsymbol{k})$具有周期性，同时也具有与晶格相同的对称性，因而 $\varepsilon(\boldsymbol{k}) = \varepsilon(-\boldsymbol{k})$。从式（3-78）可知，$\boldsymbol{k}$ 和$-\boldsymbol{k}$ 态，$\varepsilon(\boldsymbol{k})$函数具有数值相等符号相反的斜率，因此 $\boldsymbol{v}(\boldsymbol{k}) = -\boldsymbol{v}(-\boldsymbol{k})$，如图 3-31b 所示。这个关系式是讨论导体、绝缘体和半导体能带的一个重要公式。

最终根据能量梯度求得的一维情况下的电子速度如图 3-32 所示。在能带低，k 很小，晶格散射对电子运动的影响很小，$v = \hbar k/m$，与自由电子类似（图 3-32b）。随着 k 的增加，晶格散射增强，当 k 增大到布里渊区边界时，电子遭受到强布喇格反射，散射波增强到与入射波相同（但传播方向相反），这时电子速度为零。

在 $\varepsilon(k)\sim k$ 曲线的拐点处（图 3-32a 中的 c 点）电子能量最大。

3.5.2　有效质量 m^* 与波矢大小 k 的关系

本小节讨论晶体电子对外力 F 的响应。当外力等于零时，晶体电子还受到晶格势场的作用，它对电子运动的影响已概括到能量函数 $\varepsilon(k)$中，其值可通过解薛定谔波动方程求得。在有外力 F 作用时，电子除了和外力作用外，还和晶格相互作用着。电子的加速度实际上是外力 F 和晶格内部相互作用力 F_l 综合作用的结果，即 $a = \mathrm{d}v/\mathrm{d}t = (F+F_l)/m^*$。讨论电子对外力的响应时，可引入已概括了电子与晶格相互作用的有效质量 m^*来描述，把电子在外力作用下的加速度同外力 F 的关系写成类似牛顿第二定律的形式，即

$$\frac{\mathrm{d}v}{\mathrm{d}t} = \frac{F}{m^*}，\quad m^* = m\frac{F}{F+F_l} \tag{3-88}$$

因此，晶体电子对外力的响应就像具有质量为 m^*的自由电子。

m^*的数值可从 $\varepsilon(k)$求得。由于等能面的各向异性，沿 $\boldsymbol{k}$ 空间不同的方向，电子的有效质量是不同的。类比自由电子的情况，在一维情况下，认为 m^*与 $\varepsilon(k)$的关系仍是式（3-85），只不过将 m 换成 m^*而已，然后方程再对 k 求导，得

$$\frac{1}{m^*} = \frac{1}{\hbar^2}\frac{\mathrm{d}^2\varepsilon}{\mathrm{d}k^2} \tag{3-89}$$

m^*随波矢 k 变化的定性曲线如图 3-32c 所示。

不同的能带，电子的有效质量不同。在同一能带里，有效质量 m^*反比于 $\varepsilon(k)$曲线斜率的变化率 $\dfrac{\mathrm{d}^2\varepsilon}{\mathrm{d}k^2}$。图 3-32 中，在能带的下半部，由于电子的运动速度随波矢的线性增加，使有效质量保持为一正的常数，相当于电子在外力作用下作加速运动。在 c 点，$\dfrac{\mathrm{d}^2\varepsilon}{\mathrm{d}k^2}=0$，$m^*\to\infty$，相当于外力 F 对晶体电子不产生任何效应。在能带的上半部（c 点以后），由于 $\dfrac{\mathrm{d}\varepsilon}{\mathrm{d}k}$随 k 增加而减小，加速度变为负值，在外力方向保持不变的情况下，这个结果等效于有效质量的符号从正变为负。加速度为负值也说明了当晶格传递给电子的动量大于外力传递给电子的动量时，电子能够克服外力的影响而作负的加速运动。

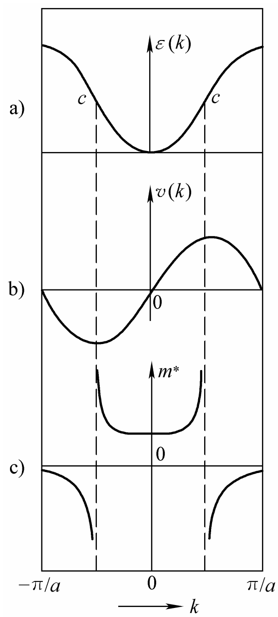


图 3-32　一维能带中的电子速度和有效质量

a）能带　b）速度　c）有效质量

有效质量的表达式直接反映了能带的曲率，也

就是说能带的宽窄与有效质量有直接的关系。所以用有效质量的概念不但能够形象地像讨论自由电子一样来讨论晶体中电子的运动，而且由于有效质量与能带结构有关，因而有助于对能带结构的研究。

3.6　能带理论的简单应用

前面简要地阐述了能带论的基本概念。利用能带论的结果可以说明固体的许多性质，这里扼要地介绍能带理论对金属、半导体和绝缘体的区分，以及用金属及合金的某些性质等来说明能带理论的成功之处和其局限性。

3.6.1　导体、半导体和绝缘体的能带

物体按导电性可分为三类：导体，其电阻率是 10^{-8}～10^{-4}Ω·m；半导体，其电阻率是 10^{-5}～10^{-7}Ω·m；绝缘体，其电阻率是 10^{8}～10^{18}Ω·m。对导电概念可以从不同的角度来认识它，经典的导电概念是指导体中的自由电子在外场作用下作定向运动；从费米分布来说，导电是指电子态在外场作用下，使费米分布出现不对称性；从能带结构来说，导电是指一部分电子获得能量后在能带中运动。能带理论能够成功地说明为什么有些晶体是导体，有些却是绝缘体和半导体。也正是以此为起点，逐步发展为导体、半导体和绝缘体的现代理论。

在无外电场下，对于在任何能带（满带和不满带）中，波矢为$\pm\boldsymbol{k}$的两个状态有相同的能量，大小相等且方向相反的电子速度。当没有外电场作用时，在一个完全被电子填满的能带中，尽管就每个电子来说有电流$-e\boldsymbol{v}$（$\boldsymbol{v}$ 电子速度），但是在 $\boldsymbol{k}$ 和$-\boldsymbol{k}$ 的电子电流$-e\boldsymbol{v}(\boldsymbol{k})$和$-e\boldsymbol{v}(-\boldsymbol{k})$正好抵消，故总的电流为零。

当存在外电场时，满带和不满带的情况不同。对于满带，在电场作用下，所有电子都以同样的速度沿电场相反的方向从一个状态变到另一个状态。尽管所有电子都在作同方向运动，但从布里渊区一边出去的电子，又在另一边同时填进来，电子运动不改变布里渊区内电子的分布情况。因此，满带不起导电作用，如图 3-33a 所示。对于不满带，只有部分状态被电子占据，电场的作用使电子的状态发生变化。由于电子在运动过程中受到晶格振动、杂质、缺陷的散射作用，最终达到一个稳定的不对称分布，如图 3-33b 所示。沿电场方向与逆电场方向电子的运动数目不相等，总电流不等于零。所以，在外场作用下不满带可以导电。对导体和非导体，有如图 3-34 中 a、b、c、d 的能带类型。在非导体中电子恰好填满最低的一系列能带，高处一系列能带全部空着，禁带宽度大，由于满带不导电，故是非导体。在导体中，除去一系列满带外，还有部分被电子填充的能带（称导带），在外电场作用下，电子产生运动，起导电的作用。另外电子虽在满带中，但由于空带和满带重迭，形成了导体。半导体和非导体能带模型相同，而半导体的禁带比

非导体的禁带窄，在 2eV 以下，在满带顶附近的部分电子在热激发下，越过禁带跑到空带中去，形成了导电。能带理论说明导体和绝缘体的区别主要在于电子在不满带中，还是在满带中。

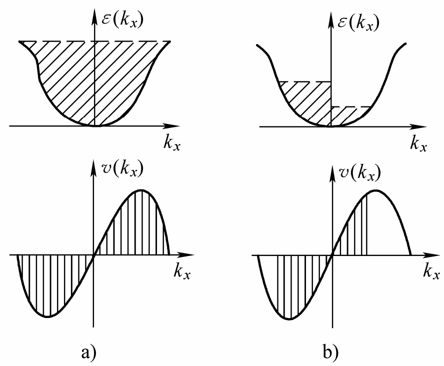


图 3-33　满带与不满带在外电场下的分布

a）满带　b）不满带

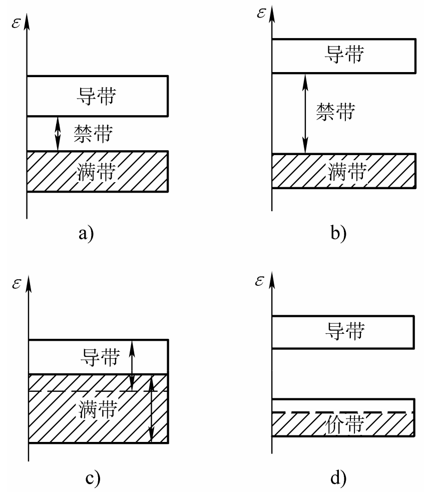


图 3-34　能带基本类型

a）半导体　b）绝缘体　c）金属（能带交叠）　d）金属（能带不交叠）

由于实际上往往发生能带的重迭，原子能级和能带之间常常不是简单地一一对应，只有对实际能带结构进行具体了解，才能分析填充能带的情况，不能以一般原理作出固体是导体还是非导体的结论。

上面的讨论对某些晶体是适合的，但对于有些晶体如氧化镍能带论并不适用，氧化镍是立方晶体，每个镍离子（Ni^{2+}）有 8 个电子，整个晶体有 8N 个电子（晶

体镍离子数），在具有立方对称的势场中，原来 5 度简并（不计自旋）的 3d 态分成一个 2 度简并和另一个 3 度简并的能带，计及自旋共可容纳 10N 个电子，现在只有 8N 个电子，所以能带不满，依照能带论应为导体，但实际上它是绝缘体。要说明这个事实，必须设想 3d 电子并不能自由运动，电子共有化模型和准自由电子近似对氧化镍不适用。虽然能带论能解释很多问题，但由于其本身是一种近似，具有一定的局限性，对有些晶体并不适用。

最后讨论“空穴”的概念，设想在满带中某一状态 $\boldsymbol{k}$ 未被电子占据，此时能带中空出一个位置，则能带不满，在电场作用下应有电流产生，用 $\boldsymbol{j}_k$ 表示。如果再引入一个电子填补这个空状态，这个电子的电流为 $-e\boldsymbol{v}(\boldsymbol{k})$，引入这个电子后，能带又被充满，总电流仍为零。所以 $\boldsymbol{j}_k+[-e\boldsymbol{v}(\boldsymbol{k})]=0$，即

$$\boldsymbol{j}_k = e\boldsymbol{v}(\boldsymbol{k}) \tag{3-90}$$

式（3-90）表示当 $\boldsymbol{k}$ 态空着时，能带中的电流犹如一个带正电荷 e 所产生的电流，而其运动速度等于在 $\boldsymbol{k}$ 态的电子速度 $\boldsymbol{v}(\boldsymbol{k})$，这个空的状态称为“空穴”。在电磁场作用下，“空穴”的位置变化（指 $\boldsymbol{k}$ 空间的状态变化）和周围电子的能态变化是一样的，如同前进队伍中缺少了一个人，这个空位可随队伍一起前进。空状态 $\boldsymbol{k}$ 变化的规律为

$$\frac{\mathrm{d}\boldsymbol{k}}{\mathrm{d}t} = \frac{1}{h}\left[-e\boldsymbol{E} - e\left(\frac{1}{h}\nabla_{\boldsymbol{k}}\varepsilon \times \boldsymbol{B}\right)\right] \tag{3-91}$$

式中，$\boldsymbol{E}$ 和 $\boldsymbol{B}$ 分别为电场强度和磁感应强度。由于满带顶的电子比较容易被激发到导带，因此“空穴”多位于能带顶。在能带顶附近电子的有效质量 $m^*<0$，即在能带顶的电子加速度犹如一个具有质量 $m^*=-m_h$ 的粒子。故

$$\frac{\mathrm{d}\boldsymbol{v}}{\mathrm{d}t} = -\frac{1}{m_h}\left[-e\boldsymbol{E} - e\left(\frac{1}{h}\nabla_{\boldsymbol{k}}\varepsilon \times \boldsymbol{B}\right)\right] = \frac{1}{m_h}\left[e\boldsymbol{E} + e\left(\frac{1}{h}\nabla_{\boldsymbol{k}}\varepsilon \times \boldsymbol{B}\right)\right] \tag{3-92}$$

上式犹如一个具有正电荷 e，正质量 m_h 的粒子在电磁场中运动所产生的加速度。因此“空穴”的运动规律和一个带正电荷 e、正质量 m_h 的粒子的运动规律完全一样。

3.6.2　金属的能带

周期表上的元素，分为金属和非金属两类。金属元素又可分为简单金属和过渡金属两类。凡是内电子壳层完全被填满或完全空着的元素均属于简单金属，如 K、Ca 的 3d 壳层完全空着，而 Cu 的 3d 壳层完全填满，故它们属于简单金属。内电子壳层未完全填满的元素属过渡金属，如 Cr、Ni、V 等 3d 壳层都未填满，属过渡金属。这两种金属元素中，外层电子较少，易失去电子而形成正离子，故称正电性金属，它富有金属性。金属性较弱的元素，如 As、Sn、Sb、Bi 等也称半金属。

碱金属其原子都有一个 s 价电子，如 Na 为 $3s^1$ 电子。这些原子组成晶体时，它们的 s 能级分裂成很宽的 s 能带，而且是半填满，其能带结构如图 3-35a 所示。在 3s 价带上，只有一半的能级被电子所占据。自然，这些被电子占据的能级应该是能量较低的能级，而 3s 价带中能量较高的处于上方的能级很少有电子占据。当温度为绝对零度时，只有下面一半的能级被电子占据，上面一半的能级没有电子占据。能带中占据的最高能级称为费密能级。而当温度大于绝对零度时，有一些电子获得了能量，跳到价带里的较高能级，而在相对应的较低的能级上失去了电子，产生了相同数量的空穴，如图 3-36 所示。这些激发电子和空穴都是携带电荷的载流子。这些金属在室温下是体心立方结构，碱金属中的传导电子有效质量如表 3-1 所示。可见 Na 的传导电子的有效质量 m^*接近于自由电子的质量 m，因此 Na 的传导电子行为在某些方面和理想自由电子气相像。

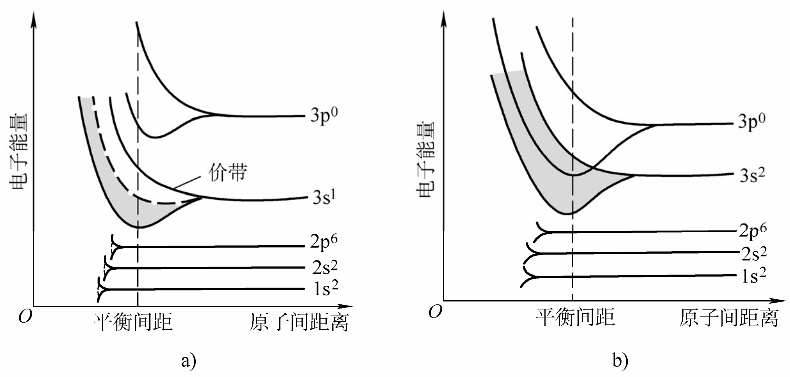


图 3-35　能带

a）Na 的能带　b）Mg 的能带

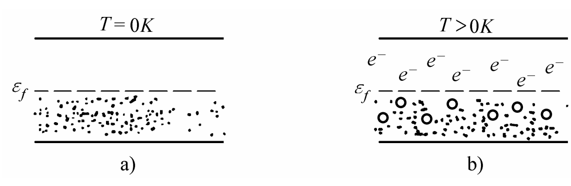


图 3-36　能带中电子随温度升高而进行能级跃迁

a）绝对零度时，所有外层电子占据尽可能低的能级

b）温度升高时，部分电子被激发到原来未被填充的能级

表 3-1　碱金属和贵金属的电子有效质量

元素	碱金属					贵金属		
	Li	Na	K	Rb	Cs	Cu	Ag	Au
m^*/m	1.40	0.98	0.94	0.87	0.83	1.022	0.994	0.994

图 3-35b 表示碱土金属镁的能带结构。像镁这样的周期表ⅡA 族元素的最外层 3s 轨道有 2 个电子，所以按理说它的 3s 能带就会被电子全部占满。但是，由于固体镁的 3p 能带与 3s 能带有重叠，这种重叠使得电子能够激发到 3s 和 3p 的重叠能带里的高能级，所以镁具有导电性。铝和其他的周期表ⅢA 族元素的 3s 能带和 3p 能带也有类似的能带重叠现象。

在从钪到镍的过渡族金属中，未被电子充满的 3d 能带和 4s 能带发生重叠。这种重叠使得电子能够被激发到高能量的能级。能带之间的复杂的相互作用使得这些金属的导电性不够理想。但铜是一个例外。铜中的内层 3d 能带已经被电子充满，这些电子被原子紧紧束缚，不能与 4s 能带相互作用。由于铜中的 3d 能带和 4s 能带之间基本没有相互作用，所以铜的导电性非常好。银和金的情况与铜类似。贵金属的电子有效质量与自由电子的质量都相接近，如表 3-1 所示。

在表 3-2 中列出了过渡族金属 Ni、Pd 和 Pt 与相邻贵金属 Cu、Ag、Au 各自的结合能和它们的原子外层电子结构。从表中可见，过渡族金属 Ni、Pd、Pt 的结合能比其相邻的贵金属 Cu、Ag、Au 的结合能大得多。上述特点可定性地由晶体的能带理论加以解释。可以假定过渡族金属与相邻贵金属的能带形状大致相同，其结合能的差别可由电子在能带中的填充情况来解释。d 壳层的半径比外面 s 价电子所处壳层半径小很多。当金属原子互相靠近形成晶体时，d 壳层的电子云相互重叠较少，而 s 价电子壳层重叠特别多。因此 s 带的特点是很宽，能量上限高，可容纳的电子少（共可容纳 2N 个，N 为组成晶体的原子数）；相比之下，d 带又低又窄，可以容纳的电子数多（共可容纳 10N 个电子），能级密度大（s 带比 d 带能级密度小许多）。图 3-37 画出了 Fe、Ni、Co 的 3d、4s 能带的 $a(\varepsilon)$曲线。图中虚线表示电子填充情况。因为 3d 和 4s 带有交叠，故研究电子在能带中的填充情况时，要同时考虑到 3d 电子和 4s 电子的数目。图中虚线上的数字，表示一个原子填充 3d 电子和 4s 电子的数目。现以 Cu 和 Ni 为例说明。Ni 每个原子的 3d 和 4s 电子数共 10 个，3d 带没填满，4s 带和 3d 带的交叠部分填充到同样程度；Cu 每个原子的 3d 和 4s 电子数共 11 个，3d 带填满后，又填 4s 带至半满，能量高出 3d 带很多。这是因为 4s 带很宽，能带密度很小，每多填一个电子，电子能量增加很多。这样，虽然 Cu 比 Ni 只多一个电子，但其电子费密能级却比 Ni 高许多，故 Ni 的结合能比 Cu 大。总之，过渡族金属的 d 带不满，且能级低而密，可容纳较

表 3-2　过渡族金属及贵金属的结合能

元素	外层电子结构	结合能/$J\cdot mol^{-1}$	元素	外层电子结构	结合能/$J\cdot mol^{-1}$
Ni	$3d^34s^2$	3.56×10^5	Cu	$3d^{10}4s^1$	3.39×10^5
Pd	$4d^{10}$	4.16×10^5	Ag	$4d^{10}5s^2$	2.85×10^5
Pt	$5d^36s^1$	5.32×10^5	Au	$5d^{10}6s^1$	3.86×10^5

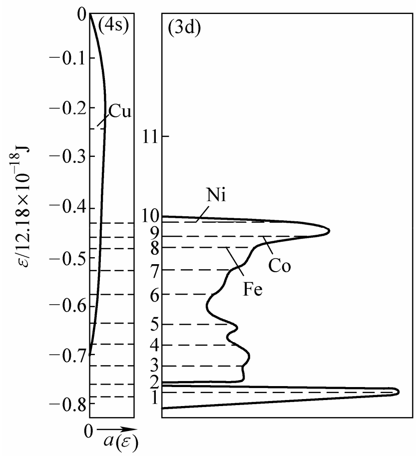


图 3-37　3d、4s 带的能级密度

多的电子，夺取较高的 s 带中的电子，降低费密能级。据测 Ni 的 s 壳层只有 0.54 个电子，而 1.46 个电子被夺到 3d 壳层中去。它们所具有的特殊物性，如结合能大、高热容、高电阻率、铁磁性及磁性反常都与其电子能带结构有关。

3.6.3　合金中的能带

金属元素可以互相溶解形成固容体或者进行化合形成化合物，另外也能形成电子化合物中间相。中间相这种化合物是由两种金属组成的，其一种为一价的金属（如 Cu、Ag、Na 等）或过渡金属（Mn、Fe、Co 等），而另一种为二至五价的金属（如 Be、Mg、Zn 等）。这种化合物的特点是，价电子数与原子数之间具有一定的比例，即 3:2、21:13、7:4 等，每一比值均对应于一定的晶格结构。例如成分 CuZn、Cu_3Al 及 Ag_3Al 的合金相，虽然它们的原子百分数很不相同，但价电子总数与原子总数之比（称电子浓度）相同，即为 1.5。它们的晶体结构都是体心立方结构。Hume Bothery 研究电子化合物得出：当电子浓度为 3:2 时，化合物具有体心立方结构（β 相）。当电子浓度为 21:13 时，具有复杂的立方晶体结构，单位晶胞有 52 个原子（γ 相）。电子浓度为 7:4 时，具有密排六方晶格结构（ε 相）。电子浓度决定电子化合物的规律可用金属电子论解释。

图 3-38 为面心立方晶格和体心立方晶格晶体的状态密度曲线。当电子所占的最高能级尚未达到面心立方晶体的第一布里渊区边界或者在布里渊区边界左右时，排布同样的电子数，面心立方晶格晶体的总能量 ε_{fcc} 小于体心立方晶格的总能量 ε_{bcc}，面心立方晶体稳定。但是随着电子数 N 的增加（超过布里渊区边界），面心立方晶体的电子密度 $a(\varepsilon)$急剧下降，而体心立方晶体的电子密度 $a(\varepsilon)$急剧上升。以致 ε_{fcc} 大于体心立方晶格的总能量 ε_{bcc}，这时晶体结构以体心立方稳定。据此可

以求出面心立方晶格，体心立方晶格以及其他晶格稳定时的最大电子数。

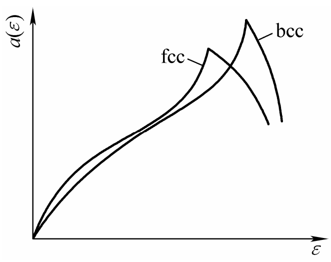


图 3-38　fcc、bcc 晶体状态密度曲线示意图

面心立方晶格的第一布里渊区为多面体（14 面体，8 个六角形面和 6 个正方形面组成），估计晶格稳定的最大电子数时，以等能面为球面近似，球面的半径是最短的满足布里渊区边界条件的波矢 $\boldsymbol{k}$。对面心立方晶格，晶格常数为 a，则 k 等于布里渊区中心到六角形面的距离，亦等于体心立方格子中心到任意一个角的距离之半（也就是体心立方对角线的 1/4）。面心立方倒格子为体心立方格子，格子边长为 $2/a$（体心立方倒格子为面心立方格子，格子边长为 a）。则

$$k=\frac{1}{4}\left(\sqrt{3}\times\frac{2}{a}\right)=\frac{\sqrt{3}}{2a} \tag{3-93}$$

设晶体中有 N 个原子，在面心晶格中单位原胞包含 4 个原子，每个原子所占的体积为 $a^3/4$。故晶体的体积 $V=Na^3/4$，亦等于 $\boldsymbol{k}$ 空间的状态密度（不计自旋），即 $V=\mathrm{d}N/(\mathrm{d}k_x\mathrm{d}k_y\mathrm{d}k_z)$，这里 $\mathrm{d}k_x$、$\mathrm{d}k_y$、$\mathrm{d}k_z$ 为 $\boldsymbol{k}$ 空间的三个分量。由于在 $\boldsymbol{k}$ 空间的状态密度为 $2V$（计入自旋）。则在球内能容纳的最大电子数 N_e 等于球体积乘以 $2V$，即

$$N_e=\frac{\pi\sqrt{3}}{2a^3}\times2\times\frac{Na^3}{4}=\frac{\sqrt{3}}{4}\pi a^3=1.36N \tag{3-94}$$

因此，在面心立方相，当每个原子平均有 1.36 个价电子时，电子正好填充到距原点最近的布里渊区边界上的能级。

同理，可以算出体心立方晶格对应等能面为球面近似的布里渊区能含的最大电子数 N_e 为

$$N_e=2\times\frac{4}{3}\pi\left(\frac{\sqrt{2}}{2a}\right)^3\times\frac{a^3N}{2}=\frac{2\sqrt{2}}{6}\pi N=1.48N \tag{3-95}$$

就是说，当每个原子平均有 1.48 个价电子时，电子正好填充到同 12 面体的面相接触的能级。

上述讨论得出金属电子论可以说明电子化合物的形成，主要决定于电子浓度的规律。仔细讨论起来，问题并不那么简单。①在一些包含过渡元素的中间相中，

过渡元素的价数如何确定，如 CoAl、NiAl、FeAl 等合金相是体心立方结构，欲归类于 3:2 的电子化合物，则 Co、Ni、Fe 等过渡元素为零价。但从另外一些合金相，如 $CrAl_7$、Co_2Al_9、$MnAl_6$、$FeAl_3$ 等发现，当过渡元素的 d 带中空穴数减少时，铝的含量也相应减少，因而设想铝的价电子填入到 d 带，过渡元素似乎具有负的价数。②布里渊区多面体，以最小 k 的球面近似计算最大电子数是建立在假定球形的等能面碰到布里渊区边界时才发生畸变。但有人测 Cu 的等能面，发现尚未碰到布里渊区边界时球形的等能面早已发生畸变。③分析表明，组元有明确的价的情况下（例如 Cu 为 1 价，Zn 为 2 价，Ga 为 3 价）溶质原子的 s+p 电子不一定全部进入导带，而有可能束缚在溶质原字上。Mott 指出：Zn、Ga、Ge 溶于 Cu 中至少贡献一个电子给导带，第二个电子或在导带或在束缚带，Ga、Ge 的第三个电子肯定在束缚带，这一点可以用组元在纯态的第二电离能佐证。Zn 的第二电离能是解离成二价离子所需的能量，为 ε_2，金属晶体的结合能为 ε_b，则

$$\varepsilon_b = -[V(r_0) + \bar{\varepsilon}_0] - \varepsilon_1 \tag{3-96}$$

式中，ε_1 为溶剂的一次离化能；$V(r_0)$是溶剂原子实与价电子的吸引势能；$\bar{\varepsilon}_0$ 是零点平均动能。如果 $\varepsilon_2 < V(r_0)$，则第二个电子可能进入导带，研究得出 Cu 的 $\varepsilon_b + \bar{\varepsilon}_0 + \varepsilon_1 = 15.35\text{eV}$。而 Zn、Ga、Ge 的第二电离能分别为 17.89eV、20.4eV、15.36eV，既然溶质原子的价电子不一定全部进入导带，合金的价电子如何计算就成为问题。

第 4 章　晶格振动和晶体的热学性质

在研究晶体时，认为各原子占据着点阵的结点，为了简化也常假设各原子在结点上是静止不动的。实际上，晶体内的原子并不是在各自的平衡位置上固定不动的，而是围绕其平衡位置作热振动。热振动的剧烈程度与温度有关。当温度不太高时，各原子在平衡位置附近作微小振动；温度升高，少数原子离开平衡位置作较长距离的运动，产生扩散现象；温度再高，绝大多数原子作剧烈运动，破坏了周期性，结果导致晶体的熔化。人们把原子在平衡位置附近的微小振动叫做晶格振动。晶格振动以波的形式在晶格中传播叫做格波。由于晶体内原子间存在着相互作用力，各个原子的振动也并非是孤立的，而是相互联系着的，因此在晶体中形成了各种模式的波。只有当振动甚为微弱时，原子间非谐的相互作用才可以忽略，即在简谐近似下，这些模式才是相互独立的。由于晶格的周期性条件，模式所取的能量值不是连续的而是分立的。对于这些独立而又分立的振动模式，可用一系列独立的简谐振子来描述。和光子的情形相似，这些谐振子的能量量子 $\hbar\omega$ 称为声子，其中 ω 是振动模式的角频率。这样晶格振动的总体就可看做是声子的系统。若原子间是非谐的相互作用可看做微扰项，则声子间发生能量交换，并且在相互作用过程中，某种频率的声子产生，另外频率的声子则湮灭。

晶格振动对晶体的许多性质有重要的影响，例如固体的热容、热膨胀、热传导等直接与晶格振动有关。这些都可以用声子的概念来描述，晶格的振动使电子受到散射而电阻增加，这可看做是电子受到声子的碰撞。另外，晶体中的光学性质也与晶格振动有密切关系，这在很大程度上可看做是光子与声子的相互作用乃至强烈的耦合。

4.1　晶格振动的经典理论

固体是由大量的原子组成的，原子又可分成离子和价电子两部分。离子与离子，离子与电子，电子与电子之间存在相互作用，这是复杂的多体问题。要在理论上解决晶格振动问题，必须作近似处理，进行必要假设近似。①绝热近似。严格说来，离子运动会引起电子云畸变，而电子运动又会影响离子的振动状态。所以离子的运动必须和电子的运动一起考虑。但是离子比电子重 10^3～10^5 倍，因此离子的运动速度比电子小得多。电子的费米速度是 $10^6 m\cdot s^{-1}$，而典型的离子速度至多是 $10^3 m\cdot s^{-1}$。现在考虑离子运动时，可近似认为电子能很快地适应离子的位置变

化，在离子运动的任何一个瞬时，电子都处于基态。这样电子对离子运动的影响，可以通过引入均匀分布的空间负电荷所产生的常量势场来近似处理，把电子运动和离子运动分开，这称为绝热近似。在绝热近似下，我们能单独处理离子运动问题，由于电子可以很快地适应离子位置的变化，于是离子的运动可以看成是中性原子的运动，没有必要在术语上区别是离子振动还是原子振动，但通常利用原子振动这个术语。②简谐近似。在温度不太高的情况下，把原子在平衡位置附近的微小振动看成是简谐振动。

晶格振动是个很复杂的问题，为了能够简单地抓住主要特点加以讨论，先考虑一维晶格的振动。然后把所得的一些主要结论和主要方法加以推广，应用到三维晶格的振动。

4.1.1　一维简单晶格振动

如图 4-1 所示的一维原子链，每个原子都具有相同的质量 m，平衡时原子间距（晶格常数）为 a。由于热运动各原子离开了它的平衡位置，用 x_n 代表第 n 个原子离开平衡位置的位移，第 n 个原子和第 n+1 个原子的相对位移为 $x_{n+1}-x_n$。不但原子链是一维的，原子的振动也是一维的。下面先求由于原子间的相互作用，原子所受到的恢复力与相对位移的关系。

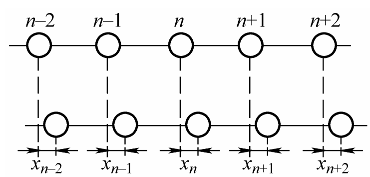


图 4-1　一维单原子链的振动

设两原子的互作用势能是 $U(r)$，r 为两原子间的间距。在平衡位置时，两个原子间的互作用势能是 $U(a)$，令 $\delta = x_{n+1}-x_n$，则产生相对位移后，相互作用势能变成 $U(a+\delta)$。将 $U(a+\delta)$在平衡位置附近用泰勒级数展开，得到

$$U(a+\delta)=U(a)+\frac{1}{1!}\left(\frac{\mathrm{d}U}{\mathrm{d}r}\right)_a\delta+\frac{1}{2!}\left(\frac{\mathrm{d}^2U}{\mathrm{d}r^2}\right)_a\delta^2+\frac{1}{3!}\left(\frac{\mathrm{d}^3U}{\mathrm{d}r^3}\right)_a\delta^3+\cdots \tag{4-1}$$

式中首项 $U(a)$为常数，次项为零（因为平衡时势能取极小值）。又因 δ 很小，势能展开式（4-1）中可只保留到 δ^2（后面的项为非简谐项）。这时两原子间的恢复力 F 为

$$F=-\frac{\mathrm{d}U(r)}{\mathrm{d}r}=-\frac{\mathrm{d}U(r)}{\mathrm{d}\delta}=-\left(\frac{\mathrm{d}^2U}{\mathrm{d}r^2}\right)_a\delta=-\beta\delta \tag{4-2}$$

式中 $\beta=\left(\frac{\mathrm{d}^2U}{\mathrm{d}r^2}\right)_a$。式（4-2）指出，当相对位移很小时，$\beta$ 为常数（称为恢复力常数），作用力为准弹性力，与 δ 成正比，此方法叫简谐近似。

第 n 个原子除受第 n–1、n+1 相邻原子作用外，还受其他原子的作用，即受整个晶体原子的作用，为了简化只考虑相邻原子的互相作用。第 n 个原子的运动方

程为

$$m\frac{\mathrm{d}^2x_n}{\mathrm{d}t^2}=\beta(x_{n+1}+x_{n-1}-2x_n)\quad(n=1, 2, 3, \cdots, N)\tag{4-3}$$

N 为晶体的原子数。对于每一个原子，都有一个类似的运动方程，因此方程数目和原子数目相同。

设式（4-3）的解是一振幅为 A，角频率为 ω 的简谐振动

$$x_n=A\mathrm{e}^{\mathrm{i}(qna-\omega t)}\tag{4-4}$$

式中，qna 表示第 n 个原子振动的位相因子，即第 n 个原子与第 0 个原子振动的位相差，显然相邻原子的位相差是 qa；$q=2\pi/\lambda$，令 $\boldsymbol{n}$ 代表沿格波传播方向的单位矢量，则 $\boldsymbol{q}=(2\pi/\lambda)\boldsymbol{n}$ 称为波矢；λ 为波长。

将式（4-4）代入式（4-3），可得

$$\omega^2=\frac{2\beta}{m}[1-\cos(qa)]\tag{4-5}$$

亦即

$$\omega=2\left(\frac{\beta}{m}\right)^{1/2}\left|\sin\left(\frac{qa}{2}\right)\right|\tag{4-6}$$

式（4-6）称为一维单原子晶格振动的色散关系式。不同的 q 对应着不同的 ω，即不同的波矢的波，波速（相速）$\upsilon_p=\omega/q$ 不同。

如果第 m 个和第 n 个原子的位相因子之差$(qma-qna)$为 2π 的整数倍，即 $ma-na=2\pi s/q$（s 为任意整数）时，

$$x_m=A\mathrm{e}^{\mathrm{i}(qma-\omega t)}=A\mathrm{e}^{\mathrm{i}(qna-\omega t)}=x_n\tag{4-7}$$

由式（4-7）可得出如下结论：①第 m 个和第 n 个原子的距离（$ma-na$）为 $2\pi/q$ 的整数倍时，原子因振动而产生的位移相等。这样可以看成 m 原子振动引起了邻近原子的振动，依次传播下去，引起 n 原子振动与 m 原子振动完全相同。这就说明原子在平衡位置附近的振动以前进波的形式在晶格中传播，这种由于晶格振动发出的波叫做格波。在晶体中格波是所有原子以相同频率，彼此之间有一定位相差的集体运动。实际上这种集体运动形成一个准连续的波。②当 $s=1$ 时，即 $ma-na=2\pi/q$。由于两振动相同的原子间距离为波长，所以格波的波长 $\lambda=2\pi/q$。③波数范围的限制——格波限制条件。由色散关系式（4-6），若 $q'=q+2\pi g/a$ 时（g 为任意整数），则

$$\omega(q')=2\left(\frac{\beta}{m}\right)^{1/2}\left|\sin\left(\frac{q'a}{2}\right)\right|=2\left(\frac{\beta}{m}\right)^{1/2}\left|\sin\left(\frac{aq}{2}\right)\right|=\omega(q)\tag{4-8}$$

由式（4-8）可知，当 q' 与 q 相差 $2\pi g/a$ 时，代表格点相同的振动状态，为了使振

动状态与 q 的单值性对应，所以把 q 限制在主值区域（$-\pi/a$，π/a），该区域称为格波的限制条件。

由色散关系式（4-6）可知，当 q 很小时，即在长波的情形下，$\sin(qa/2)\approx qa/2$，所以 $\omega=(\beta/m)^{1/2}aq$。即长波时频率与波矢是线性关系，与弹性波相似。这时波速 $v_p=a(\beta/m)^{1/2}=(E/\rho)^{1/2}$，$E$ 是弹性模量，ρ 为线密度，v_p 是常数。当 q 很大时，即在短波的情形下，由 $\sin(aq/2)$最大值为±1，ω 有最大值，$\omega_{max}=2(\beta/m)^{1/2}$，一维简单晶格的 q 与 ω 关系如图 4-2 所示。ω 随 q 呈周期性变化。$+q$ 对应前进波，$-q$ 对应与前进波方向相反的波。

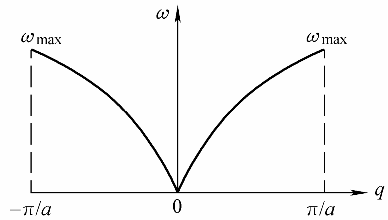


图 4-2　一维简单晶格（即布喇菲格子）的 q 与 ω 关系

原子在平衡位置的振动以波的形式在晶体中传播，一维简单晶格振动的波形如图 4-3 所示。

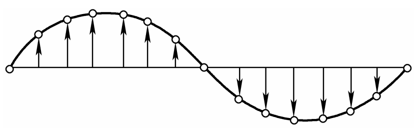


图 4-3　晶体中原子在平衡位置的振动

4.1.2　一维复式格子振动

为简单起见，考虑由两种不同种原子所构成的一维复式格子，相邻同种原子间的距离为 $2a$（$2a$ 是这复式格子的晶格常数），如图 4-4 所示。质量为 m 的原子位于 $2n-1$，$2n+1$，$2n+3$，…各点；质量为 M 的原子位于 $2n-2$，$2n$，$2n+2$，…各点。原子振动是一维的。类似于式（4-3）得到

$$\left.\begin{aligned} m\frac{\mathrm{d}^2x_{2n+1}}{\mathrm{d}t^2}&=\beta(x_{2n+2}+x_{2n}-2x_{2n+1}) \\ M\frac{\mathrm{d}^2x_{2n+2}}{\mathrm{d}t^2}&=\beta(x_{2n+3}+x_{2n+1}-2x_{2n+2}) \end{aligned}\right\} \tag{4-9}$$

为明确起见，这里假设 $M>m$，与式（4-3）相似，方程组（4-9）的解也可以是角频率为 ω 的简谐振动

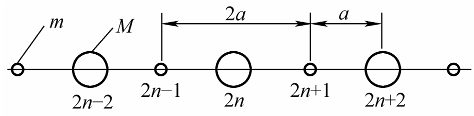


图 4-4　一维复式格子

$$\left.\begin{aligned} x_{2n+1} &= A\mathrm{e}^{\mathrm{i}[q(2n+1)a-\omega t]} \\ x_{2n+2} &= B\mathrm{e}^{\mathrm{i}[q(2n+2)a-\omega t]} \end{aligned}\right\} \tag{4-10}$$

由于这里包含有两种不同的原子，这两种不同原子振动的振幅，一般来说也是不同的。

把式（4-10）代入式（4-9），得

$$\left.\begin{aligned} -m\omega^2 A &= \beta(\mathrm{e}^{\mathrm{i}qa}+\mathrm{e}^{-\mathrm{i}qa})B-2\beta A \\ -M\omega^2 B &= \beta(\mathrm{e}^{\mathrm{i}qa}+\mathrm{e}^{-\mathrm{i}qa})A-2\beta B \end{aligned}\right\} \tag{4-11}$$

上式又可改写为

$$\left.\begin{aligned} (2\beta-m\omega^2)A-[2\beta\cos(qa)]B=0 \\ -[2\beta\cos(qa)]A+(2\beta-M\omega^2)B=0 \end{aligned}\right\} \tag{4-12}$$

若 A、B 有异于零的解，则其系数行列式必须等于零，即

$$\begin{vmatrix} 2\beta-m\omega^2 & -2\beta\cos(qa) \\ -\beta\cos(qa) & 2\beta-M\omega^2 \end{vmatrix}=0 \tag{4-13}$$

由此可以解得

$$\omega^2=\frac{\beta}{mM}\left\{(m+M)\pm[m^2+M^2+2mM\cos(2qa)]^{1/2}\right\} \tag{4-14}$$

由上式可以看到，ω 与 q 之间存在着两种不同的色散关系，即对一维简单复式格子，可以存在两种独立的格波（这一点是与前面所讨论的一维简单晶格不同，对于一维简单晶格，只能存在一种格波），这两种不同的格波各有自己的色散关系

$$\omega_1{}^2=\frac{\beta}{mM}\left\{(m+M)-[m^2+M^2+2mM\cos(2qa)]^{1/2}\right\} \tag{4-15}$$

$$\omega_2{}^2=\frac{\beta}{mM}\left\{(m+M)+[m^2+M^2+2mM\cos(2qa)]^{1/2}\right\} \tag{4-16}$$

为了便于进一步讨论这两支不同格波的性质，就需要知道 qa 取值的范围。在考虑一维的简单晶格格波的色散关系时，已取 qa 属于区间（−π，π）。同样，对于一维复式格子，为了保证波函数的单值性，则把 q 值限制在（−π/(2a)，π/(2a)），

其中 $2a$ 是这复式格子的晶格常数。

现在再看式（4-15）和式（4-16），因为 $2qa$ 限制在（$-\pi$，π），所以 ω_1 的最大值为

$$\omega_{1(\max)}=\left(\frac{\beta}{mM}\right)^{1/2}\left[(m+M)-(M-m)\right]^{1/2}=\left(\frac{2\beta}{M}\right)^{1/2} \tag{4-17}$$

而 ω_2 的最小值为

$$\omega_{2(\max)}=\left(\frac{\beta}{mM}\right)^{1/2}\left[(m+M)+(M-m)\right]^{1/2}=\left(\frac{2\beta}{m}\right)^{1/2} \tag{4-18}$$

因为 $M > m$，从而 ω_2 的最小值比 ω_1 的最大值还要大。换句话说，ω_1 一支的格波频率总比 ω_2 一支的格波频率为低。实际上，ω_2 一支的格波可以用光来激发，所以常称为光频支格波，简称为光学波。而 ω_1 一支格波则称为声频支格波，简称为声学波，命名起因于这样一个事实，当 q 很小时，该格波表现为长波长的弹性波，其频率与速度无关，频率与波矢是线性关系，与弹性波相似，而纵弹性波与声波是等同的。现在，由于高频超声波技术的发展，ω_1 一支也可以用超声波来激发了。

4.1.3 声学波和光学波

再来讨论复式格子中两支格波的色散关系。ω_1 一支的色散关系式（4-15）可改写为

$$\begin{aligned}\omega_1^{\,2}&=\frac{\beta}{mM}\left\{(m+M)-[(m+M)^2-2mM(1-\cos(2qa))]^{1/2}\right\}\\&=\frac{\beta}{mM}(m+M)\left\{1-\left[1-\frac{4mM}{(m+M)^2}\sin^2(qa)\right]^{1/2}\right\}\end{aligned} \tag{4-19}$$

如果 $\dfrac{4mM}{(m+M)^2}\sin^2(qa)\ll 1$（这在实际上是成立的），则式（4-19）近似地化为

$$\omega_1=\left(\frac{2\beta}{m+M}\right)^{1/2}\left|\sin(qa)\right| \tag{4-20}$$

把式（4-20）与式（4-6）比较，可见 ω_1 一支的色散关系与一维布喇菲格子的情形形式上是相同的，也具有如图 4-2 所示的特征。这也就是说，由完全相同原子所组成的布喇菲格子只有声学波。

ω_2 一支的色散关系式（4-16）则可改写为

$$\begin{aligned}\omega_2^{\,2}&=\frac{\beta}{mM}\left\{(m+M)+[(m+M)^2-2mM(1-\cos(2qa))]^{1/2}\right\}\\&=\frac{\beta}{mM}(m+M)\left\{1+\left[1-\frac{4mM}{(m+M)^2}\sin^2(qa)\right]^{1/2}\right\}\end{aligned} \tag{4-21}$$

在 $\frac{4mM}{(m+M)^2}\sin^2(qa)<1$ 的条件下，则式（4-21）近似地化为

$$\omega_2^{\ 2}=\frac{2\beta}{mM}\left(m+M\right)\left[1-\frac{mM}{(m+M)^2}\sin^2(qa)\right] \qquad (4\text{-}22)$$

可见，当 $q\to 0$（即波长 λ 很大）时，光学波的频率具有最大值为

$$\omega_{2(\max)}=\left(\frac{2\beta}{\mu}\right)^{1/2} \qquad (4\text{-}23)$$

其中 $\mu = mM/(m+M)$是两种原子的折合质量。而当 $q\to 0$ 时，由式（4-20）看出，$\omega_1\to 0$，这时，声学波频率则为最小。

综合本节以及式（4-17）和式（4-18）的结果，归纳如下：①声学波的频率 ω_1 最大值为$(2\beta/M)^{1/2}$（当 $q=\pm\pi/(2a)$时），最小值为 0（当 $q\to 0$ 时）。②光学波的频率 ω_2 最大值为$(2\beta/\mu)^{1/2}$（当 $q\to 0$ 时），最小值为$(2\beta/m)^{1/2}$（当 $q=\pm\pi/(2a)$时）。

一维双原子复式格子中，声学波与光学波的色散曲线如图 4-5 所示。

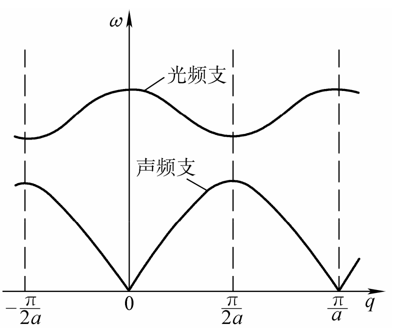


图 4-5　一维双原子复式格子的振动频谱

再看相邻两种原子振幅之比，这由式（4-12）决定。对于声学波，相邻两种原子振幅之比为

$$\left(\frac{A}{B}\right)_1=\frac{2\beta\cos(qa)}{2\beta-m\omega_1^{\ 2}} \qquad (4\text{-}24)$$

因为 $\omega_1^{\ 2}<\frac{2\beta}{M}$（$m<M$），而一般 $\cos(qa)>0$，所以$\left(\frac{A}{B}\right)_1>0$。这就说，相邻两种不同原子的振幅都有相同的正号或负号，即对于声学波，相邻原子都是沿着同一方向振动的，其振动概况如图 4-6 所示。当波长相当长时，声学波实际上代表原胞（由异质两原子组成）质心的振动。

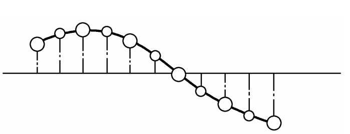


图 4-6　声学波示意图

对于光学波，相邻两种原子振幅之比为

$$\left(\frac{A}{B}\right)_2=\frac{2\beta-M\omega_2^{\ 2}}{2\beta\cos(qa)} \qquad (4\text{-}25)$$

因 $\omega_2^{\ 2}>\frac{2\beta}{m}$（$m<M$），而一般 $\cos(qa)>0$，所以$\left(\frac{A}{B}\right)_2<0$。由此可见，对于光学波，相邻两种不同原子的振动方向是相反的。而当 q 很小时，$\cos(qa)=1$，又 $\omega_2^{\ 2}=$

$\frac{2\beta}{\mu}$，得出$\left(\frac{A}{B}\right)_2=-\frac{M}{m}$。因此对于波长很长的光学波（长光学波），$mA+MB=0$，即原胞的质心保持不动。由此也可定性地看出，光学波是代表原胞中两个原子的相对振动。光学波的振动概况如图 4-7 所示。

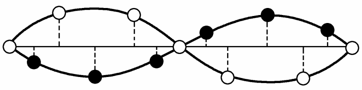


图 4-7 光学波示意图

4.1.4 周期性边界条件（玻恩—冯卡门（Born-von-Karman）边界条件）

上面由振动波函数单值的要求，考虑了波矢 q 的取值范围，对于一维布喇菲格子，q 属于（$-\pi/a,\pi/a$），对一维双原子的复式格子，q 值限制在（$-\pi/(2a),\pi/(2a)$）。但波矢的值并不是连续的，而只能在上述范围内取有限数目的分立值，这是由边界条件所决定的。

为了说明这个问题，考虑由 N 个相同原子构成的一维晶体，很显然边界上原子所处的情况与体内原子的不同，边界处原子的振动状态就该和内部原子的有所差别，在前面的讨论中，没有考虑到这个边界问题，认为一维晶体是无限的。但实际晶体总是有限的，总存在着边界，而此边界对内部原子的振动状态总会有所影响。玻恩和冯卡门把边界对内部原子振动状态的影响考虑成如下面所述的周期性边界条件。设想在一长为 $N\ a$ 的有限晶体边界之外，仍然有无穷多个相同的晶体，并且各块晶体内相对应的原子的运动情况一样，即第 j 个原子和第 $tN+j$ 个原子的运动情况一样，其中 $t=1, 2, 3, \cdots$。这样设想的无限晶体中的原子和原来实际的有限晶体中的原子，两者所受到的互作用势能却是有差别的。但是进一步的分析表明，由于互作用主要是短程的，实际的有限晶体中只有边界上极少数原子的运动才受到相邻的假想晶体的影响。就有限晶体而言，绝大部分原子的运动实际上不会受到这些假想晶体的影响。

在上述假想的周期性边界条件下，对于一维有限的布喇菲格子，第 n 个原子应和第 $n+N$ 个原子振动情况相同，即 $x_n=x_{n+N}$，再由式（4-4）可得 $e^{iqNa}=1$，即

$$qNa=2\pi l \tag{4-26}$$

式中 l 为整数。由式（4-26）可以看出，描写晶格振动状态的波矢 q 只能取一些分立值。因为 q 限制于（$-\pi/a$，π/a），所以 l 位于区间（$-N/a$，N/a）。

把 l 限制在

$$-\frac{N}{2}<l\leqslant\frac{N}{2} \tag{4-27}$$

则 q 限制在

$$-\frac{\pi}{a}<l\leqslant\frac{\pi}{a} \tag{4-28}$$

对于一维复式格子，设晶体有 N 个原胞（每一个原胞含两个不同原子），根据周期性边界条件 $x_{2n+1}=x_{2(n+N)+1}$，得到 $e^{i2qNa}=1$，即 $2qNa=2\pi l$（l 为整数）。关系式（4-27）同样适用，这里 q 限制在

$$-\frac{\pi}{2a}<l\leqslant\frac{\pi}{2a} \tag{4-29}$$

所以，一维复式格子的 q 也只能取 N 个不同的值。波矢 q 的数目亦即振动状态的数目，等于原胞数目。在波矢空间，一维双原子复式格子的每一个可能的 q 所占据的线度为 $\pi/(Na)$。这里，对应于每一个 q 值有两个不同的 ω，一个是光学波角频率，另一个是声学波角频率。因此，对于一维双原子复式格子，角频率数为 $2N$，既然每一角频率对应于一个格波，格波数必为 $2N$。在一维双原子复式格子中，每个原胞有两个原子，晶体的自由度是 $2N$，因此得到这样结论：晶格振动波矢的数目等于晶体原胞数；晶格振动频率的数目等于晶体的自由度数。这些结论对于三维晶格振动也适用。

4.2　晶格振动的量子化

晶格振动是晶体中诸原子（离子）集体在作振动，其结果表现为晶格中的格波。一般而言，格波不一定是简谐的，但可以展成为简谐平面波的线性叠加。当振动微弱时，即相当于简谐近似的情况，格波直接就是简谐波。这时，格波之间的相互作用可以忽略，从而可以认为它们的存在是相互独立的，称为独立的模式。每一独立的模式对应一个振动态。晶格的周期性又给予了格波以一定的边界条件（玻恩—冯卡门条件），使得独立的模式亦即独立的振动态是分立的。因此，我们可以用独立简谐振子的振动来表述格波的独立模式，这就是声子概念的由来。声子就是晶格振动中的简谐振子的能量量子化。

从数学运算的角度看，由于晶格振动是晶体中诸原子的集体运动，系统的总能量的表达式中亦即总的哈密顿量的式中必然包含诸原子的速度和坐标，例如，势能函数中包含有依赖于两两原子坐标的交叉项，这就带来了理论表述上的困难。但在简谐近似下，如上所述，晶体中格波的模式可以表述为稳定的独立模式。为此，引进正则坐标，通过正则变换，把原来的坐标系变换成正则坐标系，就可能消去势能中的交叉项，使得系统的哈密顿量能够表述成标准式（法式），即哈密顿量对角化了。这样，就可把晶格振动的总能量表述为独立简谐振子能量之和。下面再把以上的说明，用数学的形式表述出来。

4.2.1　一维简单晶格振动的总能量

为简单起见，以一维单原子晶体振动的量子化为例，引出声子概念，然后推

广到三维晶格的情况。

一维单原子链（参照图 4-1。N 个原子，原子质量为 m，原子间距为 a，恢复力常数为 β），在简谐近似下，只考虑最近邻相互作用，那么总势能 U 为

$$U = \frac{\beta}{2}\sum_n (x_{n+1} - x_n)^2 = \frac{\beta}{2}\sum_n (x_{n+1}^2 + x_n^2 - 2x_{n+1}x_n) \tag{4-30}$$

总动能 T 为

$$T = \frac{1}{2}\sum_n m\left(\frac{\mathrm{d}x_n}{\mathrm{d}t}\right)^2 = \frac{1}{2}\sum_n m\dot{x}_n^2 \tag{4-31}$$

系统的总能量 H 为（即哈密顿量）

$$H = U + T = \frac{1}{2}\sum_n m\dot{x}_n^2 + \frac{\beta}{2}\sum_n (x_{n+1} - x_n)^2 \tag{4-32}$$

势能式（4-30）中包括了交叉项，通过坐标变换使交叉项消除。

运动方程式（4-3）的特解为式（4-4），即 $x_n = A\mathrm{e}^{\mathrm{i}(qna-\omega t)}$，一般解应为特解的线性组合，为

$$x_n = \frac{1}{\sqrt{Nm}}\sum_q Q(q,t)\mathrm{e}^{\mathrm{i}qna} \tag{4-33}$$

在式（4-33）中，式（4-32）的时间因子和振幅都包含在系数 $Q(q, t)$中，而且质量因子也分离出来。式（4-33）实际上是代表 $x_n(t)$在 q 空间的傅里叶展开式。

式（4-33）中的 $\mathrm{e}^{\mathrm{i}qna}$，实际上代表一些独立的模式（$q$ 可取正值或负值，代表前进或后退简谐波），$\mathrm{e}^{\mathrm{i}qna}$ 具有正交性。根据 $\mathrm{e}^{\mathrm{i}qna}$ 的正交性和拉普拉斯函数，最终可得哈密顿量 H 为

$$H = \frac{1}{2}\sum_q [P(q,t)^2 + \omega(q)^2 Q(q,t)^2] \tag{4-34}$$

式中，$P(q,t) = \dfrac{dQ(q,t)}{dt} = \dot{Q}(q,t)$，$\omega(q) = \dfrac{2\beta}{m}[1 - \cos(qa)]$。

经过变换后，哈密顿量（总能量）已不包含交叉项，成为经典谐振子哈密顿量之和，新的动量和空间坐标已不是原来原子的动量和空间坐标，而是反映整体运动的坐标，它称为简正坐标。

根据经典哈密顿量式（4-34），可以直接用来作为量子力学分析的出发点，此时要把 $P(q, t)$和 $Q(q, t)$看做量子力学的正则算符，那么哈密顿量变为

$$H = \sum_q \frac{1}{2}\left[-\hbar^2 \frac{\partial^2}{\partial Q(q)^2} + \omega(q)Q(q)^2\right] \tag{4-35}$$

相应的薛定谔方程为

$$\left[\sum_{q}\frac{1}{2}\left(-\hbar^{2}\frac{\partial^{2}}{\partial Q(q)^{2}}+\omega(q)Q(q)^{2}\right)\right]\psi=E\psi \tag{4-36}$$

式（4-36）中的 E 即（4-35）的 H。式（4-36）是量子力学中的一系列独立简谐振子的薛定谔方程。其独立简谐振子的能量 ε 为

$$\varepsilon=\left(n(q)+\frac{1}{2}\right)\hbar\omega(q)\quad (n(q)=0,1,2,\cdots) \tag{4-37}$$

ε 是量子化的。一维晶格振动的总能量 E 为

$$E=\sum_{q}^{N}\varepsilon(q)=\sum_{q}^{N}\left[n(q)+\frac{1}{2}\right]\hbar\omega(q) \tag{4-38}$$

式中 N 为一维晶格的原胞数，它等于独立简谐振子数。由玻恩—冯卡门边界条件也可看出，q 只能取 N 个不同的值。

4.2.2　三维复式格子的总能量

一维晶格振动量子化的主要方法和结论可以推广到三维晶格振动系统。对于含有 N 个原胞、每个原胞包括 l 个原子的三维晶格系统，其总能量 E 为

$$E=\sum_{q}^{N}\sum_{j}^{3l}\left[n_{j}(q)+\frac{1}{2}\right]\hbar\omega_{j}(q)=\sum_{qj}^{3lN}\left[n_{j}(q)+\frac{1}{2}\right]\hbar\omega_{j}(q)\quad (n_j(q)=0,1,2,\cdots) \tag{4-39}$$

从式（4-39）可知，共有 $3lN$ 种不同的振动模式（可以把一个原子空间振动分解为空间相互垂直的三个方向的振动），即 $3lN$ 个不同的 $\omega_j(q)$值。$\hbar\omega_j(q)$是晶格振动能量的增减单位，也就是说晶格振动能量的增减必须是 $\hbar\omega_j(q)$的整数倍，这种能量单元(或称能量量子)我们称为声子。

从声子概念出发，可以重新描述晶格振动问题。晶格振动是一种整体的激发形式，晶格振动时产生了声子，声子的能量是 $\hbar\omega_j(q)$，$\omega_j(q)$是声子的频率，共有 $3lN$ 种不同的频率模式。如果振动模从基态（能量为(1/2) $\hbar\omega_j(q)$）激发到能量为 $[n_j(q)+1/2]\ \hbar\omega_j(q)$的激发态，那么就产生了能量为 $\hbar\omega_j(q)$的 $n_j(q)$个声子。共有 3 支声学声子，$3l-3$ 支光学声子，根据声子偏振方向与波矢传播方向关系，又分纵向声子和横向声子。晶格振动系统可以转变为声子系统来研究，在简谐近似中，各振动模都是相互独立的，该系统为无相互作用的声子气系统。如果考虑非谐效应，那么声子和声子之间就存在相互作用。

下面我们分几个方面讨论一下声子问题。

① 任意一谐振子其基态能量为 $\varepsilon_0=(1/2)\,\hbar\omega$，具有激发能 $\varepsilon_n=(n+1/2)\,\hbar\omega$ 的几率 P_n 是和 $\exp(-\varepsilon_n/kT)$成正比（k 为玻耳兹曼常数），由于 $\sum_{n}P_n=1$，那么几率应

由下式给出

$$P_n = \frac{\exp(-\varepsilon_n / kT)}{\sum_n \exp(-\varepsilon_n / kT)} = \frac{\mathrm{e}^{-n\hbar\omega/kT}}{\sum_n \mathrm{e}^{-n\hbar\omega/kT}} \tag{4-40}$$

式（4-40）中分子和分母的共同因子 $\mathrm{e}^{-\frac{1}{2}\hbar\omega/kT}$ 被消去了。由于 $\sum_n x^n = (1-x)^{-1}$，所以式（4-40）变为

$$P_n = \frac{\mathrm{e}^{-n\hbar\omega/kT}}{\sum_n \mathrm{e}^{-n\hbar\omega/kT}} = \mathrm{e}^{-n\hbar\omega/kT}(1-\mathrm{e}^{-\hbar\omega/kT}) \tag{4-41}$$

由式（4-41）可以求出能量 $\varepsilon_n = (n+1/2)\,\hbar\omega$ 振子的平均能量为

$$\bar{\varepsilon} = \sum_n \varepsilon_n P_n = \sum_n \left(n+\frac{1}{2}\right)\hbar\omega P_n = \frac{1}{2}\hbar\omega + \sum_n n\hbar\omega P_n \tag{4-42}$$

利用 $\sum_n nx^n = x/(1-x)^2$，最后获得

$$\bar{\varepsilon} = \frac{\hbar\omega}{2} + \frac{\hbar\omega}{\mathrm{e}^{\hbar\omega/kT}-1} = \left(\frac{1}{\mathrm{e}^{\hbar\omega/kT}-1} + \frac{1}{2}\right)\hbar\omega \tag{4-43}$$

又根据式（4-38），处于简振模 $\omega_j(q)$振子的平均能量应为

$$\bar{\varepsilon} = \left[n_j(q) + \frac{1}{2}\right]\hbar\omega_j(q) \tag{4-44}$$

比较式（4-44）和式（4-43），有

$$n_j(q) = \frac{1}{\mathrm{e}^{\hbar\omega_j(q)/kT}-1} \tag{4-45}$$

从声子概念出发，$n_j(q)$是处于 j、q 态的声子平均数。因此由式（4-45）知，声子遵循玻色统计分布，声子是玻色子。

② 在简谐近似中，声子气是无相互作用的声子构成的，声子是玻色子，因此晶格振动的每个状态能够被任何数目的不可区别的声子占据，声子数仅与晶格振动的能量值有关，即依赖于温度。在 T=0K 时，没有任何声子被激发。

正像我们可以把电磁波看成光子流一样，也可以把弹性波(在晶体中即为格波)看成声子流。和光子一样，声子是玻色子。

4.3 固体的热容

热容是分子或原子热运动的能量随温度而变化的一个物理量。热容是物体温度升高 1K 所需要增加的能量。显然，物体的质量不同，热容不同。1g 物质的热容称为“比热容”，1mol 物质的热容称为“摩尔热容”。另外，物体的热容还与

它的热过程有关，假如加热过程是恒压条件下进行的，所测定的热容称为定压热容(C_p)。假如加热过程保持物体容积不变，所测定的热容称为定容热容(C_V)。它们分别为

$$C_V=\left(\frac{\partial Q}{\partial T}\right)_V=\left(\frac{\partial E}{\partial T}\right)_V,\quad C_p=\left(\frac{\partial Q}{\partial T}\right)_p=\left(\frac{\partial H}{\partial T}\right)_p \tag{4-46}$$

式中 Q 为热量，E 为内能，H 为焓。由于恒压加热过程中，物体除温度升高外，还要对外界做功，所以温度每提高 1K 需要吸收更多的热量，即 $C_p>C_V$。对于物质的凝聚态，当压力不太大时，固体的体积变化可略去，C_p 与 C_V 的差异可以忽略。

在大多数固体中，供给晶格振动的能量是对热容的主要贡献，在非铁磁性绝缘体中它是唯一的贡献。在金属中对热容的其他贡献来自传导电子，而在磁性材料中对热容的其他贡献则来自磁有序。本节只讨论晶格振动对热容的贡献。

4.3.1 晶态固体热容的经验定律和经典理论

在 19 世纪已发现了两个有关晶体热容的经验定律。一是元素的热容定律——杜隆—珀蒂定律：恒压下元素的原子摩尔热容为 25J/(K·mol)。实际上，大部分元素的原子摩尔热容都接近该值，特别在高温时符合得更好。只是轻元素的原子摩尔热容与此值相差较大。另一个是化合物的热容定律——柯普定律：化合物分子热容等于构成此化合物各元素原子热容之和。

经典的热容理论可对此经验定律作如下解释。在固体中可以用谐振子代表每一个原子在一个自由度的振动。根据经典理论，能量按自由度均分，每一振动自由度平均动能是$(1/2)kT$（参照理想气体分子的平均动能），因为是谐振子，平均势能等于平均动能，故平均势能也为$(1/2)kT$。那么每一个自由度的平均能量为 kT。一个原子有三个振动自由度，一个原子的平均能量为 $3kT$。1mol 固体中有 N_0 个原子，总能量为 $E=3N_0kT=3RT$。这里 $N_0=6.023\times10^{23}$，为阿伏加德罗常数，T 为热力学温度，k 为玻耳兹曼常数，摩尔气体常数 $R=8.314\text{J/(K·mol)}$。根据摩尔定容热容定义

$$C_{V,\mathrm{m}}=\left(\frac{\partial E}{\partial T}\right)_V=\left[\frac{\partial(3N_0kT)}{\partial T}\right]_V=3N_0k=3R\approx25(\mathrm{J/(K\cdot mol)}) \tag{4-47}$$

在简谐近似下，原子的振动是严格简谐的，不会发生热膨胀现象，因而定容的前提条件也成立。

由式（4-47）可知，摩尔定容热容是与温度无关的常数，这就是杜隆—珀蒂定律。对于双原子的固态化合物，1mol 中的原子数为 $2N_0$，故摩尔定容热容为 $C_{V,\mathrm{m}}=2\times25\text{J/(K·mol)}$，三原子固态化合物的摩尔定容热容 $C_{V,\mathrm{m}}=3\times25\text{J/(K·mol)}$，余此类推。杜隆—珀蒂定律在高温时与实验结果是很符合的。但在低温时，摩尔定容热容的实验值并不是一个恒量。它随温度降低而减小，在接近 0K 时，摩尔定容热容值按

T^3 的规律趋于零（绝缘体）或按 $AT+BT^3$ 的规律趋于零（金属）。如图4-8为Cu的 $C_{V,\mathrm{m}}$ 与 T 关系曲线（$C_{V,\mathrm{m}}$ 值不是直接测量的，而是由测量 $C_{p,\mathrm{m}}$ 值换算来的），曲线可分三个区，每区的特点：Ⅰ区 $C_{V,\mathrm{m}}\propto T$；Ⅱ区 $C_{V,\mathrm{m}}\propto T^3$；Ⅲ区逐渐过渡为平缓上升的直线。对于低温下摩尔定容热容减小的现象使经典理论遇到了困难，需要用量子理论来解释。

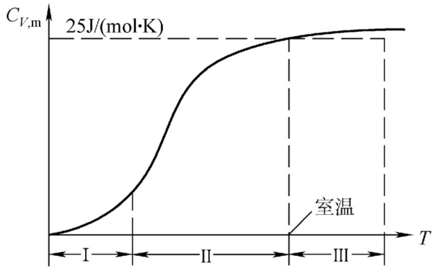


图4-8　Cu的 $C_{V,\mathrm{m}}$ 随 T 变化曲线

4.3.2　晶态固体热容的量子理论

根据晶格振动的量子化概念导出的的三维晶格系统的总能量 E 的公式（4-39）及量子平均数公式（4-45），可以导出该晶体的摩尔定容热容 $C_{V,\mathrm{m}}$（N 个原胞，每个原胞有个 l 原子）为

$$C_{V,\mathrm{m}}=\left(\frac{\partial E}{\partial T}\right)_V=\sum_{qj}^{3Nl}k\left(\frac{\hbar\omega_j(q)}{kT}\right)^2\frac{\mathrm{e}^{\hbar\omega_j(q)/kT}}{\left(\mathrm{e}^{\hbar\omega_j(q)/kT}-1\right)^2} \tag{4-48}$$

这就是按照量子理论求得的摩尔定容热容表达式。但是由上式计算 $C_{V,\mathrm{m}}$ 必须知道谐振子的频谱。这是非常困难的事。实际上，采用简化的爱因斯坦模型和德拜模型。

1. 爱因斯坦模型

爱因斯坦提出的假设是：每一个原子都是一个独立的振子，振子的能量是量子化的。振动产生的声子为玻色子。原子之间彼此无关，并且都是以相同的角频率振动。根据式（4-48），而 $\omega_j(q)=\omega$（j=1, 2, 3, …。每个振子的频率相同），$Nl=N_A$（总原子数），因此晶体的摩尔定容热容为

$$C_{V,\mathrm{m}}=\left(\frac{\partial E}{\partial T}\right)_V=3N_Ak\left(\frac{\hbar\omega}{kT}\right)^2\frac{\mathrm{e}^{\hbar\omega/kT}}{\left(\mathrm{e}^{\hbar\omega/kT}-1\right)^2} \tag{4-49}$$

定义 $\Theta_E=\dfrac{\hbar\omega}{k}$ 为爱因斯坦温度。求晶体的摩尔定容热容时，$N_A=N_0$，$N_0k=R$。晶体的摩尔定容热容 $C_{V,\mathrm{m}}$ 为

$$C_{V,\mathrm{m}}=3R\left(\frac{\Theta_E}{T}\right)^2\frac{\mathrm{e}^{\Theta_E/T}}{\left(\mathrm{e}^{\Theta_E/T}-1\right)^2} \tag{4-50}$$

式（4-50）称为爱因斯坦量子摩尔热容公式，它表明的摩尔热容与温度的关系和实验结果有更一步的接近。大多数固体的 Θ_E 值在 100～300K 的范围内。Θ_E 可由实验确定，即在 100～300K 范围内，对应一个温度测出一个 $C_{p,\mathrm{m}}$ 值，再换算成 $C_{V,\mathrm{m}}$ 值代入式（4-50），即可求出 Θ_E。许多简单固体的摩尔热容大体符合式（4-50），所以可以认为这个理论在一级近似上描述了晶体的摩尔热容。例如，金刚石的实验数据与爱因斯坦理论曲线符合得很好。

从爱因斯坦摩尔热容模型式（4-50）可以得出

① 当高温时，$T \gg \Theta_E$，则 $\Theta_E/T \ll 1$，$\mathrm{e}^{\Theta_E/T} \approx 1+\dfrac{\Theta_E}{T}$。所以

$$C_{V,\mathrm{m}} = 3R\left(\frac{\Theta_E}{T}\right)^2 \frac{\mathrm{e}^{\Theta_E/T}}{\left(\mathrm{e}^{\Theta_E/T}-1\right)^2} = 3R\left(\frac{\Theta_E}{T}\right)^2 \frac{1+\dfrac{\Theta_E}{T}}{\left(1+\dfrac{\Theta_E}{T}-1\right)^2} = 3R \qquad (4\text{-}51\mathrm{a})$$

这就是杜隆—珀蒂定律，也就是实验曲线图 4-8 的Ⅲ区的情况。说明在高温时，爱因斯坦模型同经典热容理论和实验结果符合得很好。这是很好理解的，因为在高温区，振子的能量近似等于 kT，而 kT 远大于能量量子($\hbar\omega$)时，量子化效应就可以忽略，所以给出这个结果与高温区的热容实验结果符合。

② 当低温时，$T \ll \Theta_E$，则 $\mathrm{e}^{\Theta_E/T} \gg 1$，所以

$$C_{V,\mathrm{m}} = 3R\left(\frac{\Theta_E}{T}\right)^2 \mathrm{e}^{-\Theta_E/T} \qquad (4\text{-}51\mathrm{b})$$

由上式可见，$C_{V,\mathrm{m}}$ 与 T 的关系的趋势是符合实验情况的（图 4-8 的Ⅰ区），如图 4-9 所示，这是经典理论所不能得到的结果。但 $C_{V,\mathrm{m}}$ 与 T 呈指数关系变化，而试验的 $C_{V,\mathrm{m}}$ 与 T 呈三次方的关系变化。

③ 在温度中温区域（图 4-8 的Ⅱ区），爱因斯坦模型比实验值下降的快，如图 4-9 所示。

爱因斯坦模型的结果，在有些区域与实验不符。其原因是假设与实际情况不符。实际晶体质点间是互相联系的，这种联系在低温下表现得尤其显著，格波振动的频率也是有差别的。

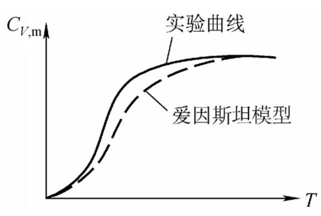


图 4-9　爱因斯坦模型与实验曲线的比较

2. 德拜热容模型

德拜考虑晶体中点阵格点间的相互作用。其主要观点是：晶格的每个谐振子不是相互独立的，而是互相联系的，把格波看成弹性波（不考虑光学波）。每个谐振子的频率不同，频率范围从零到某一最大值 ω_m。

由式（4-48）可知，一个振子对摩尔定容热容的贡献为

$$C_{V,\mathrm{m}}^{1}=k\left(\frac{\hbar\omega}{kT}\right)^{2}\frac{\mathrm{e}^{\hbar\omega/kT}}{\left(\mathrm{e}^{\hbar\omega/kT}-1\right)^{2}} \tag{4-52}$$

现在考虑到 ω 不同，若 $\rho(\omega)\mathrm{d}\omega$ 为 1mol 原子晶体频率位于 ω 与 $\omega+\mathrm{d}\omega$ 之间的振子数，则所有振子对热容的贡献就变成对式（4-52）积分的形式

$$C_{V,\mathrm{m}}=\int_{0}^{\omega_m}k\left(\frac{\hbar\omega}{kT}\right)^{2}\frac{\mathrm{e}^{\hbar\omega/kT}}{\left(\mathrm{e}^{\hbar\omega/kT}-1\right)^{2}}\rho(\omega)\,\mathrm{d}\omega \tag{4-53}$$

定义 $\Theta_D=\dfrac{\hbar\omega_m}{k}$，称为德拜温度。由此得到德拜热容模型

$$C_{V,\mathrm{m}}=3R\left[12\left(\frac{T}{\Theta_D}\right)^{3}\int_{0}^{\Theta_D/T}\frac{\chi^{3}\mathrm{d}\chi}{\mathrm{e}^{\chi}-1}-\frac{3\left(\dfrac{\Theta_D}{T}\right)}{\mathrm{e}^{\Theta_D/T}-1}\right] \tag{4-54}$$

式中 $\chi=\dfrac{\hbar\omega}{kT}$。$\Theta_D$ 与 Θ_E 类似，物质不同 Θ_D 不同，如 Na、Ti、Al、金刚石分别为、158K、420K、428K、2230K。

从德拜热容模型式（4-54）可以得出：

① 当高温时，$T\gg\Theta_D$，则 $\mathrm{e}^{\chi}\approx 1+\chi$，所以

$$C_{V,\mathrm{m}}=3R\left[12\left(\frac{T}{\Theta_D}\right)^{3}\int_{0}^{\Theta_D/T}\frac{\chi^{3}\mathrm{d}\chi}{\mathrm{e}^{\chi}-1}-\frac{3\left(\dfrac{\Theta_D}{T}\right)}{\mathrm{e}^{\Theta_D/T}-1}\right]=3R\left[12\left(\frac{\Theta_D}{T}\right)^{3}\frac{\left(\dfrac{\Theta_D}{T}\right)^{3}}{3}-3\right]=3R \tag{4-55}$$

这就是杜隆—珀蒂定律，也就是实验曲线图 4-8 的III区的情况。说明在高温时，德拜热容模型同经典热容理论和实验结果符合。

② 当低温时，$T\ll\Theta_D$，则

$$C_{V,\mathrm{m}}=\frac{12}{5}\pi^{4}R\left(\frac{T}{\Theta_D}\right)^{3} \tag{4-56}$$

对一定的物质有一定的 Θ_D，所以，$C_{V,\mathrm{m}}$ 与 T^3 成正比，它与实验符合。这比爱因斯坦理论更进一步了。

③ 在温度中温区域（图 4-8 的 II 区），德拜模型的值与实验结果也符合得很好。这也说明在一般温度下，晶体温度升高时所吸收的热量，主要来加强晶格振动。

上面所述的符合仍是大致上的，将德拜模型的值与实验结果严格对照如图 4-10 所示。在图 4-8 的 I 区，$T<5\mathrm{K}$，实验值 $C_{V,\mathrm{m}}$ 与 T 成正比；在III区，实验

值 $C_{V,\mathrm{m}}$ 虽然很接近 25J/(mol·K)，但并不是以 3R 为渐近线，而是超过 3R 继续有所上升。其原因是德拜热容模型只考虑了晶格振动对热容的贡献，但实际上，自由电子对热容也有部分贡献，热容的电子贡献部分在极高温度（对普通金属要几万度）及近于 0K 的极低温度下才是显著的。

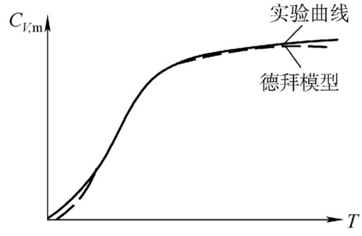


图 4-10　德拜热容模型与实验曲线的严格对比

德拜温度 Θ_D 与晶体的热容、原子间的结合力、晶体的熔点、弹性模量、温度因子都有关系。它与低温摩尔定容热容的关系如上面导出的式（4-56）。式（4-5）中，原子间的恢复力系数 β（反映原子间的结合力）与简谐振动角频率 ω^2 成正比，爱因斯坦温度 Θ_E 与 ω 成正比。一般情况下，$\Theta_E = 0.7\Theta_D$，因而 β 正比于 $\Theta_D{}^2$；一般认为在熔点 T_s 时，原子的振幅达到使晶格破坏的数值，这样原子振动的最大频率 ω_m 和熔点 T_s 间有林德曼（Lindlman）公式：$\omega_m = 2.8\times10^{12}[T_s/(mV^{2/3})]^{1/2}$，$m$ 为原子质量，V 为原子体积。由林德曼公式可以导出：$\Theta_D = 137[T_s/(mV^{2/3})]^{1/2}$；晶体的弹性模量 E 正比于 β，当然也正比于 $\Theta_D{}^2$。

3. 德拜热容模型与爱因斯坦模型本质区别及晶格热容与电子热容的关系

温度愈低，德拜热容模型近似愈好。因为在非常低的温度下，只有长波格波的激发是主要的。对于长波，晶格是可以看做连续介质的。

现在可以说明爱因斯坦模型在低温下与实验符合不好的原因。这个模型主要适用于光学波，光学波的 ω 随 q 变化很小，可以假设 ω 为常量。但是，低温下，光学波的量子化能量大于 kT，因此没有被激发。这个模型忽略了非常低频的声学波，由于低频声学波量子化能量很小，在非常低温下，它们也能吸收热量而被激发，对摩尔热容有贡献。

尽管德拜模型是相当成功的，但是，它把各种可能格波的色散关系都用连续介质中弹性波的色散关系代替，显然会有局限性。从实验发现在很宽的温度范围内 Θ_D 本身是与 T 有关的，其变化量达 10%，有时更高。为了改进德拜模型，人们应该用正确的格波色散关系代替长波近似式。

下面在德拜模型近似下，讨论晶格摩尔热容和电子摩尔热容的关系。

在晶格摩尔热容理论中的德拜温度 Θ_D 和在金属电子摩尔热容理论中的费米温度 T_f 都是经典过程和量子过程的分界限。在前一个理论中 $T > \Theta_D$ 适于用经典统计力学规律，$T < \Theta_D$ 需要用量子统计规律。在后一个理论中，$T > T_f$ 可以用经典统计规律，$T < T_f$，须要用量子统计规律。但是对于电子情况，T_f 在 10^4～10^5K 量级，所以实际温度总是远低于 T_f 的；而对晶格振动情况 Θ_D 在 10^2K 量级，实际温度往往处在量子与经典重叠区内。

在金属摩尔热容测量实验中，需要知道在什么温度下晶格摩尔热容起主导作

用（T^3 规律）。为了回答这个问题，把电子对摩尔热容的贡献 $C_{V,\mathrm{m(Z)}}^{e}$（1mol 离子提供 Z mol 自由电子）除以声子对热容的贡献 $C_{V,\mathrm{m}}^{i}$ 得

$$\frac{C_{V,\mathrm{m(Z)}}^{e}}{C_{V,\mathrm{m}}^{i}}=\frac{5}{24\pi^2}Z\frac{\Theta_D^3}{T^2T_f} \tag{4-57}$$

当 $C_{V,\mathrm{m(Z)}}^{e}=C_{V,\mathrm{m}}^{i}$ 时，$T=T_0=0.145\left(Z\Theta_D/T_f\right)^{1/2}\Theta_D$。$T>T_0$ 时，声子对摩尔热容的贡献超过电子的贡献。由于一般情况费米温度比德拜温度高 2 个量级，所以 T_0 只有几 K，这就是金属摩尔热容与 T 的线性关系只有在极低温度下才能观察到的原因。

4.3.3 固体热容与组织及状态的关系

材料的组织、状态发生变化时，有热效应产生，这个热效应就构成材料热容的附加部分，可以通过热容的附加部分来研究材料内部的组织转变。在这里主要说的是金属与合金的热容。

1. 多型性转变

多型性转变属于一级相变（第一类转变）。第一类转变是在恒温下进行，在转变温度有热函 H 的突变，并且热容成为无限大。第二类转变是在某一温度范围内进行，H 值没有跃变，只是随温度升高增加的速度变大（主要在靠近居里点 T_c 时 H 增长），并且此时热容也达到了有限的极大值。在第一种情况中，转变的潜热可以直接由 H 与 T 的关系中得到（在相应曲线上的变化）。在第二类转变中，为了确定第二类转变的热效应，需要在变化温度区间内进行热容的内插法，并且计算由试验及内插法所得曲线的面积。

图 4-11 绘出了在 0～2000K 之间铁的摩尔定压热容。图中实线是实测值曲线，虚线是 γ-Fe 在 A_3 点以下的计算摩尔定压热容值。由图 4-11 可见，铁的摩尔定压热容随温度变化过程既有第一类转变的特征，又有第二类转变的特征。在 A_2 点铁磁性的 α-Fe 转变为顺磁性的 β-Fe，这时摩尔定压热容出现极大值，A_2 点的相变属于二级相变（第二类转变）。在 A_3、A_4 及熔点（T_s），摩尔定压热容都为无限大值，故这三个温度的相变都是一级相变。

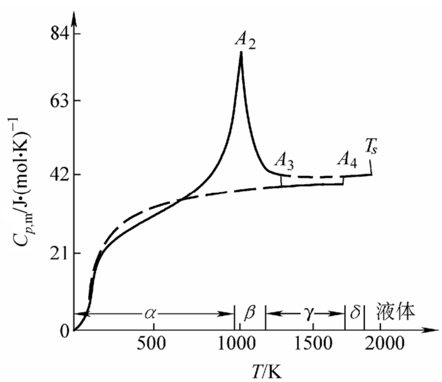


图 4-11 加热时铁的摩尔定压热容变化

2. 有序无序转变

有序无序转变是一级相变还是二级相变，目前还无定论。有人根据在相变时

热容随温度变化的特点来判别一级或二级相变。对于 Cu-Zn 合金，有序无序转变时的摩尔定压热容变化如图 4-12 所示，这里摩尔定压热容的变化有二级相变的特征，所以有人认为 Cu-Zn 有序无序转变是二级相变。

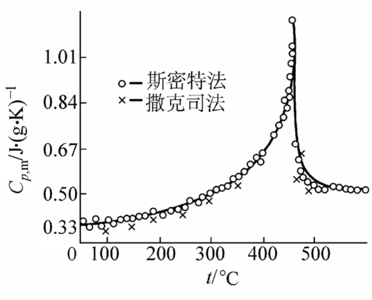


图 4-12　Cu-Zn 合金加热时摩尔定压热容变化

3. 亚稳状态的热效应

合金材料的淬火组织是典型的亚稳状态。图 4-13 是含 0.74%C 的碳钢摩尔定压热容与回火温度变化曲线。图中下部分 a）为淬火后的结果，上部分 b）是淬火回火后的结果。淬火曲线有三种不同的热效应，这三种热效应对应不同的组织结构变化。热效应Ⅰ为淬火马氏体向回火马氏体转变引起的，此时马氏体的正方度下降（即 c/a 下降），同时析出 Fe_xC 型的碳化物；热效应Ⅱ为残留奥氏体的分解引起的（淬火时，大部分奥氏体转变成马氏体，在 300℃左右分解成回火马氏体和 Fe_xC 型碳化物），分解产物是回火马氏体和 Fe_xC 型碳化物；热效应Ⅲ为马氏体的继续分解和 Fe_xC 型碳化物向 Fe_3C 的转变引起的。淬火回火曲线有两种热效应：热效应 A 是马氏体的继续分解及 Fe_xC 的析出引起的；而热效应 B 则是 Fe_xC 向 Fe_3C 的转变引起的。

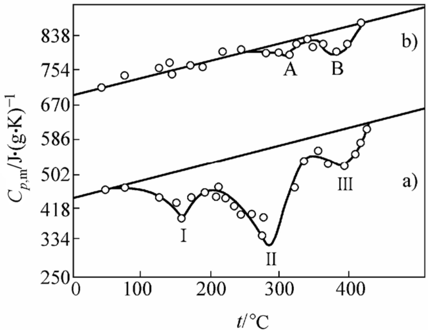


图 4-13　0.74%C 的碳钢摩尔定压热容与回火温度变化曲线

冷加工组织也是一种亚稳定状态。在金属冷加工时，一部分形变能储存于晶体中，使金属发热并使内能增加，热容及焓也增加，但是增加不多。例如加工硬化后，铁的热容比未加工前高 1.4%。金属在加工硬化时所吸收的热量在再结晶时

将放出。因此，根据加热时热容的变化，可以研究形变金属回火时的恢复和再结晶过程。

4. 化合物及多相复合材料的热容

实验证明，在较高温度下固体的摩尔定压热容大约等于构成该化合物各元素原子的摩尔定压热容之和，即

$$C_{p,\mathrm{m}} = \sum x_j C_{p,\mathrm{m}}^j \tag{4-58}$$

式中，x_j 为化合物中元素 j 原子的摩尔分数；$C_{p,\mathrm{m}}^j$ 为化合物中元素 j 原子的摩尔定压热容。这一计算式对于计算大多数氧化物和硅酸盐化合物在 573K 以上的摩尔定压热容有较好的结果。同样，对于多相复合材料有如下计算式

$$C_{p,\mathrm{m}} = \sum x_I C_{p,\mathrm{m}}^I \tag{4-59}$$

式中，x_I 为材料中第 I 组成的摩尔分数；$C_{p,\mathrm{m}}^I$ 为材料中第 I 组成的摩尔定压热容。

式（4-58）和式（4-59）有一定的普遍性，不但可以用于无机非金属材料，还可应用于金属的化合物、金属与非金属的化合物，并且还可以十分准确地用于固溶体、中间相及它们在多相合金中的混合体上。但是不适用于低温条件下，因为低温下热容与温度密切相关。另外，对于具有铁磁性的合金也不适用。

4.4 固体的热膨胀

物体的体积或长度随温度的升高而增大的现象称为热膨胀。衡量材料热膨胀现象的主要参数是膨胀系数。膨胀系数有线胀系数和体胀系数。线胀系数有某一温度区间的平均线胀系数和某一温度下的真实线胀系数之分。同样，体胀系数也有平均体胀系数和真实体胀系数之分。设 l_1、l_2 和 V_1、V_2 分别是试样在温度 T_1、T_2 下（$T_1<T_2$）的长度和体积，在温度 T_1～T_2 区间内的平均线胀系数 $\bar{\alpha}_l$ 和平均体胀系数 $\bar{\alpha}_V$ 为

$$\bar{\alpha}_l = \frac{l_2 - l_1}{l_1}\frac{1}{T_2 - T_1}，\quad \bar{\alpha}_V = \frac{V_2 - V_1}{V_1}\frac{1}{T_2 - T_1} \tag{4-60a}$$

对于立方晶系晶体，三个晶轴方向的膨胀特性相同，则 $\bar{\alpha}_V \approx 3\bar{\alpha}_l$。对于各向异性的晶体，各晶轴方向的线胀系数不同，假如分别为 $\bar{\alpha}_{l(a)}$、$\bar{\alpha}_{l(b)}$、$\bar{\alpha}_{l(c)}$，则可导出 $\bar{\alpha}_V \approx \bar{\alpha}_{l(a)} + \bar{\alpha}_{l(b)} + \bar{\alpha}_{l(c)}$。设 l 和 V 分别是试样在温度 T 下的长度和体积，那么在温度 T 下真实线胀系数 α_l 和体胀系数 α_V 为

$$\alpha_l = \frac{\mathrm{d}l}{l}\frac{1}{\mathrm{d}T}，\quad \alpha_V = \frac{\mathrm{d}V}{V}\frac{1}{\mathrm{d}T} \tag{4-60b}$$

因此，膨胀系数的定义应是当温度变化 1K 时，物质尺寸（或体积）的变化率，其

单位是 K^{-1}。

一般固体线胀系数在 10^{-2}～$10^{-6}K^{-1}$ 数量级，无机非金属材料的线胀系数一般较小，为 10^{-5}～$10^{-6}K^{-1}$ 数量级。各种金属及合金的膨胀系数相差较大，膨胀系数范围较大，在 10^{-2}～$10^{-6}K^{-1}$ 数量级。

4.4.1 晶格振动的非谐效应——热膨胀的物理本质

如果晶体中的振动是严格的简谐振动，晶体将不会因受热而膨胀。这里只以双原子分子为例，先定性地讨论热膨胀问题，所得结果可以直接应用于一维晶格，如图 4-14 所示，假定左边的原子固定不动，而右边的原子可以自由地振动。如果势能曲线对原子的平衡位置对称，则当原子振动后，其平均位置将和振幅的大小无关，如果这种振动就是热振动，则两原子间的距离将和温度无关。实际上，两原子之间的互作用势能曲线并不是严格的抛物线，而是不对称的复杂函数，如图 4-14 中所示。平衡位置的左边较陡，右边较平滑，因此当原子振动后，随着振幅（或总能量）的增加，平均位置将向右边移动。例如，当振动的总能量为某一个 E_i 时，平均位置移至 p_i。与各个能量相应的平均位置如图 4-14 中的 AB 曲线所示。物体的热膨胀就是由于势能曲线的这种不对称性所导致。

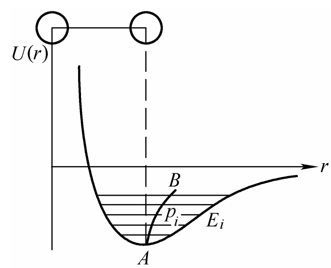


图 4-14 原子间互作用势能曲线

以下用经典的方法计算平均位置向右边移动的距离。设 r_0 是原子的平衡位置，δ 是离开平衡位置的位移。把原子在 $r_0+\delta$ 点的势能 $U(r_0+\delta)$对平衡位置 r_0 按式（4-1）展开，则

$$U(r_0+\delta)=U(r_0)+\frac{1}{1!}\left(\frac{\mathrm{d}U}{\mathrm{d}r}\right)_{r_0}\delta+\frac{1}{2!}\left(\frac{\mathrm{d}^2U}{\mathrm{d}r^2}\right)_{r_0}\delta^2+\frac{1}{3!}\left(\frac{\mathrm{d}^3U}{\mathrm{d}r^3}\right)_{r_0}\delta^3+\cdots \quad (4\text{-}61)$$

第一项为常数，第二项为零。如果取 $U(r_0) = 0$，并令式（4-61）右边第三项中 δ^2 的系数为 f，第四项中 δ^3 的系数为 $-g$，忽略 δ^3 以上各项，那么式（4-61）变成

$$U(r_0+\delta)\approx f\delta^2-g\delta^3 \quad (4\text{-}62)$$

根据玻耳兹曼统计，原子的平均位移 $\overline{\delta}$ 为

$$\overline{\delta}=\frac{\int_{-\infty}^{\infty}\delta \mathrm{e}^{-U/kT}\mathrm{d}\delta}{\int_{-\infty}^{\infty}\mathrm{e}^{-U/kT}\mathrm{d}\delta} \quad (4\text{-}63)$$

式中 k 为玻耳兹曼常数。

如果在势能的展开式中只保留 δ^2 项，即假定力是准弹性的，振动是简谐振动，则 $\bar{\delta}=0$，即原子的平均位置和平衡位置相同，没有热膨胀现象发生。如果计入非对称项，则 $\bar{\delta}\neq 0$，设 δ 很小，则式（4-63）的分子可写成

$$\begin{aligned}\int_{-\infty}^{\infty}\delta \mathrm{e}^{-U/kT}\mathrm{d}\delta &\approx \int_{-\infty}^{\infty}\delta \mathrm{e}^{-(f\delta^2-g\delta^2)/kT}\mathrm{d}\delta \approx \int_{-\infty}^{\infty}\delta \mathrm{e}^{-f\delta^2/kT}\left(1+\frac{g\delta^3}{kT}\right)\mathrm{d}\delta \\ &=2\int_{0}^{\infty}\delta \mathrm{e}^{-f\delta^2/kT}\left(\frac{g\delta^4}{kT}\right)\mathrm{d}\delta=\frac{g}{kT}\left(\frac{3}{4}\pi^{1/2}\right)\left(\frac{kT}{f}\right)\end{aligned} \tag{4-64}$$

同时，式（4-63）的分母可写成

$$\int_{-\infty}^{\infty}\mathrm{e}^{-U/kT}\mathrm{d}\delta=\int_{-\infty}^{\infty}\mathrm{e}^{-f\delta^2/kT}\mathrm{d}\delta=\left(\frac{\pi kT}{f}\right)^{1/2} \tag{4-65}$$

将式（4-64）和式（4-65）代入式（4-63）得

$$\bar{\delta}=\frac{3}{4}\frac{g}{f^2}kT \tag{4-66}$$

因此得到线胀系数 α 为

$$\alpha=\frac{1}{r_0}\frac{\mathrm{d}\bar{\delta}}{\mathrm{d}T}=\frac{3}{4}\frac{kg}{f^2 r_0} \tag{4-67}$$

显然，α 是一个与温度无关的常数。如果计入 $U(r_0+\delta)$展开式中的更高次项，则线胀系数 α 将和温度有关。

4.4.2 热膨胀系数与其他物理量之间的关系

1. 热膨胀系数与热容、温度的关系

热膨胀是固体材料受热以后晶格振动加剧而引起的体积膨胀，而晶格振动的激化就是热运动能量的增大。升高单位温度时能量的增量也就是热容的定义。所以热膨胀系数显然与热容密切相关并有着相似的规律。格留涅申（Gruneisen）从晶格振动理论推导出金属热膨胀系数与热容间的关系式（称为格留涅申关系式）为

$$\alpha_V=\frac{r}{K\nu}C_V \tag{4-68}$$

式中，α_V 为体胀系数；C_V 为质量定容热容（也叫比定容热容）；K 是体积模量；r 为格留涅申常数，表示原子非线性振动物理量，对一般物质在 1.5 与 2.5 之间变化；ν 是比体积。

从热容理论可知，在低温下 C_V 随温度的变化遵从 T^3 关系，所以在低温下膨

胀系数随温度的变化也应服从 T^3 规律，即膨胀系数与温度关系曲线同热容与温度关系曲线特征一致。实验结果也是如此，如图 4-15 所示，图中 T_s 为熔点。

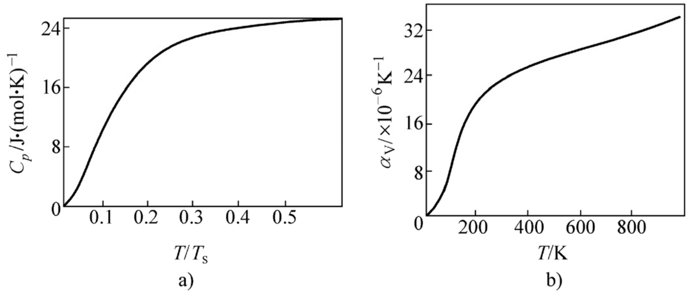


图 4-15　Al 的摩尔定压热容与温度的关系（图 a）及体胀系数与温度的关系（图 b）

由式（4-68）还可以看出，膨胀系数与比体积成反比。钢铁材料的 γ-Fe 的比体积比 α-Fe 的小，因而加热时 γ-Fe 膨胀比 α-Fe 强烈。

2. 热膨胀系数与结合能、熔点的关系

由于固体材料的热膨胀与晶体点阵中质点的位能性质有关。质点的位能性质是由质点间的结合力特性所决定的。质点间结合力很强，则位阱深而狭窄，升高同样的温度差，质点振幅增加得较少，故平均位置的位移量增加得较少，因此，热膨胀系数较小，如表 4-1 所示。

表 4-1　热膨胀系数与结合能熔点的关系

单质材料	结合能/×10^{3}J·mol^{-1}	熔点/℃	α_l/×10^{-6} K^{-1}（平均值）
金刚石	712.3	3500	2.5
硅	364.5	1415	3.5
锡	301.7	232	5.3

格留涅申还提出了关于固态的体热膨胀极限方程。方程表明，一般纯金属温度由 0K 加热到熔点 T_s，体胀为 6%。当固体加热体积增大 6%时，晶体原子之间的结合力已经很弱，因而使固态熔化变为液态。同时还可看出，物体熔点愈低，则该物质的膨胀系数愈大；反之，物体的熔点愈高，则该物体的膨胀系数愈小。这是因为在温度 0K 到熔点温度 T_s 区间，固体体积的变化总量都是 6%之故。元素的平均线胀系数 $\bar{\alpha}_l$ 与熔点 T_s 间的经验公式为

$$\bar{\alpha}_l T_s^n = A \tag{4-69}$$

式中 n=1.17，对于金属 A=7.24×10^{-2}。从该式可以看出，元素的膨胀系数愈小，则它的熔点就愈高。

3. 热膨胀系数与德拜温度的关系

前面已经叙述，由林德曼公式可以导出 Θ_D 与熔点 T_s 的关系：$\Theta_D = 137[T_s/(mV^{2/3})]^{1/2}$，若式（4-69）中 $n = 1$，那么可得到热膨胀系数与德拜温度的简单关系

$$\bar{\alpha}_l = \frac{A'}{V^{2/3}m}\frac{1}{\Theta_D^2} \tag{4-70}$$

因为熔点 T_s 和德拜温度 Θ_D 表征晶体原子间结合力（结合能），而从式（4-69）、式（4-70）不难看出了膨胀系数与原子间结合力成反比。因此，表征原子间结合力的物理量如 E、T_s、Θ_D 都与热膨胀系数有关。固体的弹性模量、熔点、德拜温度愈高，则热膨胀系数愈低。

4. 热膨胀系数与热导率的关系

弗兰克尔确定了非金属的热膨胀系数平方应正比于热阻率 ω'，即

$$\frac{1}{\lambda} = \omega' \propto C\alpha_l^{\,2} \tag{4-71}$$

式中，λ 是热导率；C 是常数。

4.4.3 固体材料热膨胀的一些规律

1. 单相材料

元素的热膨胀系数与与元素的原子序数有关，在周期表中呈周期性变化，碱金属的膨胀系数最高，过渡族金属的膨胀系数最低，如图 4-16 所示。这是由于碱金属的原子间结合力小，熔点低，所以有高的膨胀系数。

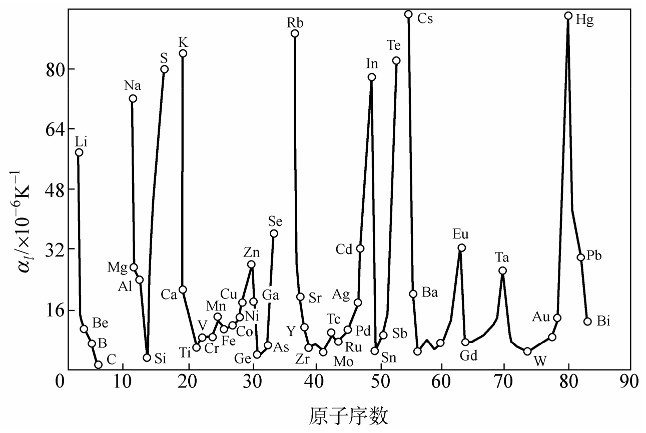


图 4-16　元素膨胀系数与原子序数之间的关系

对绝大多数合金来说，如合金内部组成是均一的单相固溶体，合金中溶质元

素的性质及含量对合金的热膨胀性能影响极为明显。一般来说，合金的膨胀系数介于组元的膨胀系数之间，符合相加律的规律。二元合金的成分和膨胀系数的关系呈一条光滑曲线，但其值比直线规律略低些。若金属固溶体基体中加入过渡族元素，则固溶体的膨胀系数变化无常。

固溶体的有序—无序转变以及其他多性型转变也会引起膨胀系数明显变化。一般磁性转变和有序—无序转变所引起的热膨胀变化特性属于二级相变，其膨胀系数与温度曲线特征是在转变点有膨胀系数的折点。金属和合金中的其他多型性转变（例如 Fe 的 $\alpha \to \gamma$ 相变）属于一级相变，它将伴随热容的突变，相应的膨胀系数将有不连续的变化，在转变点膨胀系数无限大。图 4-17 是一级、二级相变的膨胀曲线特征示意图。

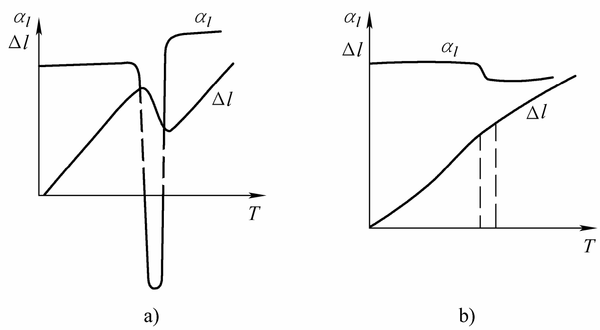


图 4-17　相变的膨胀量（Δl）$_l$ 随温度 T 变化（图 a）及线胀系数 α_l 随温度 T 变化示意图（图 b）

2. 多相复合材料

假如有一复合材料，所有组成都是各向同性的，而且均匀分布，但是由于各组成的膨胀系数不同，各组成分别都存在着内应力（拉应力或压应力）

$$\sigma_i = K_i(\overline{\alpha}_V - \alpha_i)\Delta T \qquad (4\text{-}72)$$

式中，σ_i 是第 i 部分的应力；$\overline{\alpha}_V$ 为复合体的平均体胀系数；α_i 是第 i 部分组成的体胀系数；ΔT 是从应力松弛状态算起的温度变化；$K_i = E_i/[3(1\text{-}2\mu_i)]$，$E_i$ 是第 i 部分组成的弹性模量，μ_i 是第 i 部分组成的泊松比。

整体的内应力之和为零，即

$$\sum \sigma_i V_i = \sum K_i(\overline{\alpha}_V - \alpha_i)V_i\Delta T = 0 \qquad (4\text{-}73)$$

式中，V_i 是第 i 部分组成的体积。

设：g_i 是第 i 部分组成的质量，g 是复合体的质量，$g_i/g = w_i$ 为第 i 部分组成的质量分数；ρ_i 为第 i 部分组成的密度，则 $V_i = g_i/\rho_i = gw_i/\rho_i$。将 $V_i = gw_i/\rho_i$ 代入式（4-73）得

$$\bar{\alpha}_V = \frac{\sum \alpha_i K_i w_i / \rho_i}{\sum K_i w_i / \rho_i} \tag{4-74}$$

以上把微观的内应力都看成是纯拉应力和压应力，对交界面上的切应力略而不计。假如要计入切应力的影响，情况就要复杂得多。对于仅为二相材料的情况有如下的近似式

$$\bar{\alpha}_V = \alpha_1 + V_2(\alpha_2 - \alpha_1)\frac{K_1(3K_2 + 4G_1)^2 + (K_2 - K_1)(16G_1^2 + 12G_1K_2)}{(4G_1 + 3K_2)[4V_2G_1(K_2 - K_1) + 3K_1K_2 + 4G_1K_1]} \tag{4-75}$$

式中，G_1 是第 1 部分的剪切弹性模量。

在很多情况下，式（4-74）和式（4-75）与实验结果是比较符合的。

4.5 固体热膨胀的反常现象

绝大多数金属和合金的真实膨胀系数 α 随温度 T 变化的规律，如图 4-15 所示。随温度升高，α 值先是很快地增加（$\alpha \propto T^3$），以后增加速度减慢以至近于恒值。这种情况称为正常热膨胀。但对于某些铁磁性的金属和合金，如镍和一系列铁镍合金，膨胀系数随温度的变化不符合上述规律，出现如图 4-18 的现象，即在正常热膨胀曲线上出现附加的膨胀峰，这些变化有时是非常急剧的称为反常热膨胀。镍的热膨胀峰为正，称为正反常；而 35Ni-Fe（Ni 原子的摩尔分数，%）的热膨胀峰为负，称为负反常。具有负反常膨胀特性的合金，膨胀系数可低到接近零值（甚至可达负值），或在一定温度范围内膨胀系数基本不变，故有重大的实用意义。

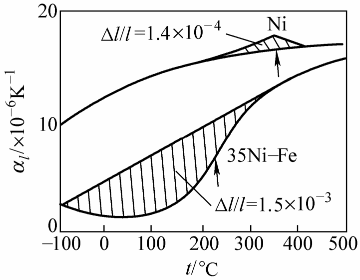


图 4-18 Ni 和 35Ni-Fe 合金的线胀系数 α 随温度 t 变化的曲线（图上虚线表示正常热膨胀时 α 值，阴影区表示反常热膨胀的范围和大小，Ni 为正反常热膨胀，35Ni-Fe 为负反常热膨胀，箭头指居里点，$\Delta l/l$ 代表最大的反常线胀系数）

反常热膨胀现象最早是 1897 年吉罗姆（Guillaume）在具有面心立方晶型的 Ni35-Fe 合金中发现的，该合金在室温附近线胀系数很小（$\alpha = 1.2\times10^{-6}\ \mathrm{K}^{-1}$）。后来在具有和铁镍合金状态图类似的 Fe-Pt、Fe-Pd 二元和三元 Fe-Ni-Co、Fe-Co-Cr

系合金中以及其他晶系和非晶态、反铁磁、亚铁磁的材料中都发现有这种反常热膨胀现象。人们把固体材料的这种膨胀系数很小或趋于零或为负值的现象叫做热膨胀反常，又称因瓦反常（Invar），它是英文 Invariation（不变）的缩写。除此之外，人们还把与此相关连的其他物理特性的反常行为也统称因瓦效应。

4.5.1　因瓦效应的微观机制

人们对因瓦效应的微观机制进行了深入研究，提出了许多的物理模型，但无一种模型占绝对优势。一般认为，因瓦效应与产生磁有序而引起体积效应是密切相关的。因此，目前因瓦理论同铁磁性理论一样分为两大支，即定域电子模型和巡回（巡游）电子模型。

1. 定域电子模型

根据定域电子模型，因瓦效应的具体机制主要有以下几种解释。

（1）交换—体积模型

图 4-19 表示不同金属的贝色—施拉特（Bethe- Slater）交换作用—原子间距曲线（图中 r_{3d} 为 3d 壳层的半径）。由图可见，在正交换作用的上升段（即曲线峰值的左侧），若原子间距为 a_1，相应的交换积分为 A_1，如交换积分增加到 A_2，相应的原子间距增加到 a_2。交换积分的增加，使得铁磁体的交换能降低。但另一方面原子间距增加发生弹性变形，提高了弹性能。当交换能与弹性能之和最小时达到平衡。当温度自居里点 T_c 以上降到 T_c 以下时发生自发磁化，由于交换积分的增加导致原子间距的改变，且是各向同性的，即发生了体积的变化。这里得到的是正的自发体积磁致伸缩，即 $\omega_s>0$（ω_s 为自发体积磁致伸缩量）。同样，在正交换作用的下降段（即曲线峰值的右侧），由于磁有序将得到负的自发体积磁致伸缩，即 $\omega_s<0$。因瓦合金的 ω_s 是一个反常大的正值，其数值比一般磁有序物质的高 1～2 个数量级。因瓦合金在贝色—施拉特曲线峰值左侧，当温度升高时，自发磁化减弱，交换积分 A 下降，导致原子间距 a 缩小，从而补偿了晶格振动引起的热膨胀，这就是因瓦合金膨胀系数很小或为负值的原因，也就是因瓦合金负反常热膨胀的物理本质。

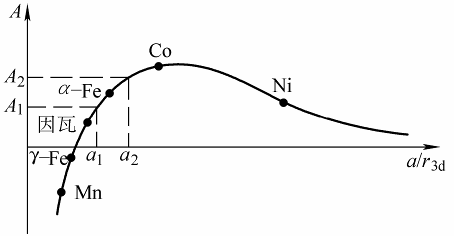


图 4-19　贝色—施拉特（Bethe- Slater）曲线

（2）潜存反铁磁性模型

康多尔斯基（Kondorsky）等人认为，在因瓦合金的铁磁基体（体积较大）中有部分的反铁磁性团（体积较小）潜存。温度升高，反铁磁性团增多，导致体积减小。例如 Fe-Ni 因瓦合金中，Ni-Ni 和 Fe-Ni 原子对呈铁磁性；Fe-Fe 原子对呈反铁磁性。

（3）双态模型

外斯（Weiss）假设因瓦合金中存在着 γ_1 和 γ_2 两种电子态。γ_1 态电子结构为 $3d^{4.0}\uparrow 3d^{3.5}\downarrow 4s^{0.5}$，是反铁磁态，有较小的体积（晶格常数为 0.354nm）和磁矩（原子磁矩为 $0.5\mu_B$）；γ_2 态电子结构为 $3d^{5.0}\uparrow 3d^{2.2}\downarrow 4s^{0.8}$，为铁磁态，有较大体积（晶格常数为 0.364nm）和磁矩（原子磁矩为 $2.8\mu_B$）。在因瓦成分区域，γ_1 和 γ_2 之间能量差小（约为 0.035eV），状态转变对外部条件很敏感。随温度升高，一些 γ_2 态热激发为 γ_1 态，由此产生大的自发体积磁致伸缩，因而抵消点阵的热膨胀。

（4）两种交换作用模型

两种交换作用模型，类似潜存反铁磁性模型，不过认为铁磁性来自 s 电子与 d 电子的 s-d 间接交换作用，反铁磁性来自 d 电子与 d 电子的 d-d 直接交换作用。温度升高时，磁矩降低，反铁磁态增加，产生抵消点阵热膨胀的热收缩效应。

（5）施罗斯（Schlosser）模型

施罗斯认为从组织结构上看合金是一种不均匀固溶体。合金由三个区域构成：核心部分是 $FeNi_3$ 型短程有序区；最外层是 γ-Fe；中间是过渡区。因瓦反常与过渡区的状态和数量有关。过渡区中的铁原子取决于近邻原子情况而表现出不同的电子状态和体积。过渡层体积愈大，因瓦反常愈显著。

（6）浓度起伏模型

卡奇（Kachi）和阿桑诺（Asano）基于低镍 Fe-Ni 合金中铁磁相和反铁磁相共存这一事实，提出了浓度起伏模型。指出，在含镍 30%～35%（原子的摩尔分数）的 Fe-Ni 合金中，相应的铁磁相比例为 50%～70%，居里温度分布在 100～500K 区间。温度升高，浓度不同，居里点也不同的铁磁相连续地从大比体积的铁磁态转变为小比体积的顺磁态，体积反常收缩。

以上各种模型均能够解释一些因瓦现象，但是由于模型中假设太多，显得这些基于定域电子模型的因瓦理论过于人为化了。而且，所有上述因瓦理论又都是以下面的条件为前题，即含铁面心立方结构和原子的（或磁性的）不均匀性。因此可以说，定域电子模型是有它的局限性的。

2. 巡游电子模型

基于巡游电子模型的斯通纳（Stoner）能带理论，许多人对因瓦问题进行了深入的研究。具体机制可分为以下几种。

（1）能带模型

认为交换作用使电子极化，并使不同极化方向的能带移位，电子动能和能态密度增大，能态宽度减小，这样导致点阵常数和体积增大，产生自发体积磁致伸缩。

（2）巡游电子弱铁磁性理论

巡游电子弱铁磁性材料都具有低居里点和低磁矩或不稳定磁有序的共同特点。由于这类材料的交换劈裂很小，可把自由能展开成级数，从而可以解释因瓦材料中的许多磁性反常和热膨胀系数很小的现象。

除上面介绍的几种因瓦理论模型外，也还有其他的理论模型。

4.5.2　因瓦合金的一些特性

因瓦合金有一些反常的物理特性。在因瓦成分范围内饱合磁化强度 M_s 较低；$\dfrac{\partial M_s}{\partial p}$ 及 $\dfrac{\partial T_c}{\partial p}$ 均为很大的负值（p 为压力）；强制体积磁致伸缩 $\dfrac{\partial \omega_s}{\partial H}$ 是很大的正值；M_s 随温度升高（直到居里点）而急剧下降；很高的强磁场下的磁化率 χ_s；以及电阻、低温热容、弹性模量、超声波吸收等的反常行为。

4.6　固体的热传导

不同的固体材料在导热性能上可以有很大的差别。有些材料是极为优良的绝热材料，有些又会是热的良导体。材料的导热性能与材料的结构和微观组织有很大关系。

4.6.1　固体材料热传导的宏观规律

当固体材料一端的温度比另一端高时，热量会从热端自动地传向冷端，这种现象就称为热传导。假如固体材料垂直于 x 轴方向的截面积为 ΔS，材料沿 x 轴方向的温度梯度为 $\dfrac{\mathrm{d}T}{\mathrm{d}x}$，在 Δt 时间内沿 x 轴正方向传过 ΔS 截面上的热量为 ΔQ，则实验表明，对于各向同性的物质，在稳定传热状态下具有如下的关系式

$$\Delta Q = -\lambda \frac{\mathrm{d}T}{\mathrm{d}x} \Delta S \Delta t \tag{4-76}$$

式中的常数 λ 称为热导率。式中负号表示热流是沿温度梯度向下的方向流动，即 $\dfrac{\mathrm{d}T}{\mathrm{d}x} < 0$ 时，$\Delta Q > 0$，热量沿 x 轴正方向传递；$\dfrac{\mathrm{d}T}{\mathrm{d}x} > 0$ 时，$\Delta Q < 0$，热量沿 x 轴负方向传递。

热导率 λ 的物理意义是指单位温度梯度下，单位时间内通过单位垂直面积的热量，所以它的单位为 $\mathrm{W/(m^2 \cdot K)}$ 或 $\mathrm{J/(m^2 \cdot s \cdot K)}$。

式（4-76）也称为傅里叶定律。它只适用于稳定传热的条件。即传热过程中，材料在 x 方向上各处的温度 T 是恒定的，与时间无关，$\dfrac{\Delta Q}{\Delta t}$ 是常数。

假如是不稳定传热过程，即物体内各处的温度随时间而变化。例如一个与外界无热交换，本身存在温度梯度的物体，随着时间的推移温度梯度趋于零的过程，就存在热端温度不断降低和冷端温度不断升高，最终达到一致的平衡温度。该物体内单位面积上温度随时间的变化率为

$$\frac{\partial T}{\partial t}=\frac{\lambda}{\rho c_p}\frac{\partial^2 T}{\partial x^2} \tag{4-77}$$

式中，ρ 为密度；c_p 为质量定压热容。

4.6.2 固体传热的微观机理

1. 声子热导

如果不考虑电子对热传导的贡献，则晶体中的热传导主要依靠声子来完成。在没有温度梯度时，声子在各方向上作无规运动。存在温度梯度时，由声子把能量从一端传递到了另一端，能流密度不为零。声子不是畅通无阻地由一端行进到另一端，在通过样品的扩散过程中遭到频繁的碰撞。通常把声子的前后两次碰撞间走过的平均距离 l 称为声子的平均自由程。利用上述模型，把经典气体分子运动论中关于热导率的理论移植到声子热传导的问题中来。声子的热导率为

$$\lambda=\frac{1}{3}c_V\bar{v}l \tag{4-78}$$

式中，c_V 是晶格（声子）质量定容热容；$\bar{v}$ 是声子平均速度的大小，就是布里渊区中全部占据态（声子）平均速度的大小。

在式（4-78）中，c_V 是与温度相关的，平均自由程 l 由散射过程决定，显然也与温度密切相关，只有 $\bar{v}$ 基本上与温度无关，所以热导率 λ 是温度的函数。c_V 与温度的关系在前面已经给出，要了解热导率 λ 的温度相关性，现在关键是了解平均自由程 l 与温度的关系。

平均自由程 l 与温度 T 的关系决定于在晶体中发生的碰撞（散射）过程，这个问题是非常复杂的。可简单归纳为如下几个重要的散射机制：声子之间的散射；声子受晶体中点缺陷（杂质、空位等）的散射；声子受样品边界的散射。与每一种散射机制相联系有一个平均自由程 l_a、l_b、l_c，因而产生相应的热阻为 W_a、W_b、W_c。总热阻为 $W=W_a+W_b+W_c$。总的平均自由程 l 为

$$\frac{1}{l}=\frac{1}{l_a}+\frac{1}{l_b}+\frac{1}{l_c} \tag{4-79}$$

如果上述几种散射强度相当，l 与 T 的关系问题必然十分复杂。这里只讨论简单情况，即某一温度范围内，假设只有一种散射机制起主导作用。

（1）声子与声子的相互作用

如果原子间的力是纯谐和的，则不存在不同声子间的碰撞机制，而平均自由程将只受到声子同晶体边界的碰撞以及点阵缺陷的限制。在有些情况下，这些效应起主导作用。然而，如果存在非谐点阵相互作用，在不同声子间就出现一种耦合，它限制平均自由程的值。非谐系统精确的简正模式不再像纯声子。下面考虑由声子相互作用引起的热阻率。

1）正常过程和倒逆过程　随着温度升高，原子间的位移变大，必须考虑非谐效应。声子不再是互相独立的，它们之间互相散射。散射过程遵守能量守恒和准动量守恒定律。设两个声子的频率和波矢分别为 ω_1、$\boldsymbol{q}_1$ 和 ω_2、$\boldsymbol{q}_2$，碰撞产生的第三个声子频率和波矢为 ω_3、$\boldsymbol{q}_3$。能量守恒：$\hbar\omega_1+\hbar\omega_2=\hbar\omega_3$；准动量守恒：$\hbar\boldsymbol{q}_1+\hbar\boldsymbol{q}_2=\hbar\boldsymbol{q}_3+\hbar\boldsymbol{G}$，这里 $\boldsymbol{G}$ 为倒格矢。

如果 $\boldsymbol{G}=\boldsymbol{0}$，$\boldsymbol{q}_1$、$\boldsymbol{q}_2$、$\boldsymbol{q}_3$ 都在第一布里渊区，声子碰撞前后系统的动量严格相等，能量不变。这样的过程不产生热阻，它不影响整个声子系统的流动。这种过程叫正常过程，也叫 N 过程。在 N 过程中，由散射引起的波矢的改变 $\boldsymbol{q}_1-\boldsymbol{q}_2-\boldsymbol{q}_3=\boldsymbol{0}$，声子集团的动量 $\boldsymbol{P}=\sum n(\boldsymbol{q})\hbar\boldsymbol{q}$ 守恒，式中 $n(\boldsymbol{q})$表示波矢为 $\boldsymbol{q}$ 的声子数。对于 $\boldsymbol{P}\neq\boldsymbol{0}$ 的分布，N 过程使 $\boldsymbol{P}$ 保持不变。即如果沿一根棒状样品推动一个初始 $\boldsymbol{P}\neq\boldsymbol{0}$ 的声子分布集团，则该分布沿这根棒传播而 $\boldsymbol{P}$ 不变。因此没有热阻，即热导率是无限的。

如果 $\boldsymbol{G}\neq\boldsymbol{0}$，这相当于 $\boldsymbol{q}_1$ 和 $\boldsymbol{q}_2$ 在第一布里渊区，$\boldsymbol{q}_3$ 不在第一布里渊区（超出第一布里渊区的范围），如图 4-20 所示。有物理意义的波矢都在第一布里渊区内。我们可以把 $\boldsymbol{q}_3$ 约化到第一布里渊区之内，用 $\boldsymbol{q}_4$ 代表，$\boldsymbol{q}_4=\boldsymbol{q}_3-\boldsymbol{G}$。值得注意的是，由声子碰撞产生的有意义的波矢 $\boldsymbol{q}_4$ 的方向与原来 $\boldsymbol{q}_1$、$\boldsymbol{q}_2$ 的方向几乎相反。在碰撞过程中初态和终态动量差非零倒格矢，这种过程对于改变声子动量起很重要的作用，是高温情况下影响平均自由程的决定性因素，这个过程引起热阻。把这种过程称为倒逆过程，也叫 U 过程。

2）倒逆过程对热导的影响　如果两个声子散射后产生第三个声子，只有 $\boldsymbol{q}_1$、$\boldsymbol{q}_2$ 波矢大小具有(1/2)$\boldsymbol{G}$ 的量级，才可能发生 $\boldsymbol{q}_1+\boldsymbol{q}_2=\boldsymbol{q}_3+\boldsymbol{G}$ 的倒逆过程，引起热阻。根据德拜理论，这两个参与碰撞的声子的能量为$(1/2)k\Theta_D$ 量级。

在高温下（$T\gg\Theta_D$），有

$$n_j(q)=\frac{1}{\mathrm{e}^{\hbar\omega_j(q)/kT}-1}\approx\frac{kT}{\hbar\omega_j(q)} \qquad (4\text{-}80)$$

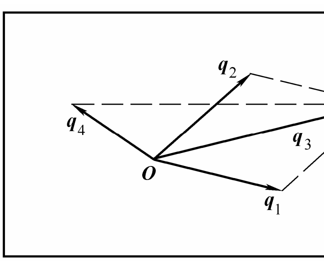


图 4-20　倒逆过程

所以总的声子数目是和 T 成正比的，能够参与倒逆过程的声子数目也与 T 成正比。由于声子数多少与碰撞几率成正比，因此，$1/\tau \propto T$（τ 为声子两次碰撞的平均时间间隔），即 $l \propto 1/T$。由于在高温下质量定容热容 c_V 遵循杜隆—珀蒂定律，与温度无关，所以此时的热导率随温度升高而降低，即

$$\lambda \propto \frac{1}{T} \tag{4-81}$$

实际上，更精确的理论分析可以导出 $\lambda \propto 1/T^x$，这里 x 介于 1 到 2 之间，这一点已经被实验所证实。

在低温下($T < \Theta_D$)，因为能够产生热阻的倒逆过程的声子能量必须为(1/2)$\boldsymbol{G}$ 量级，这样的声子数为

$$n_j(q) = \frac{1}{e^{\hbar\omega_j(q)/kT} - 1} \approx \frac{1}{e^{\Theta_D/2T} - 1} \approx e^{-\frac{\Theta_D}{2T}} \tag{4-82}$$

由此可见，温度下降时，能够参与倒逆过程的声子数按指数方式随 T 急剧减少，没有倒逆过程，热导率应是无限大的。因此可以预见，在 $T < \Theta_D$ 时，热导率公式（4-78）中的有效平均自由程将随 $e^{\Theta_D/2T}$ 改变。这种指数形式，与实验符合得很好。注意：式（4-78）中出现的声子平均自由程是声子之间倒逆过程的平均自由程，而不是声子之间所有碰撞的平均自由程。表 4-2 中列出几个代表性的平均自由程数值。

表 4-2 声子平均自由程（代表性的声速取 5×10^4m/s）

晶体	t/℃	C / J·(m^3·K)$^{-1}$	λ / W·(m·K)$^{-1}$	l / nm
石英	0	2.00×10^6	13	4
	−190	0.55×10^6	50	54
NaCl	0	1.88×10^6	7	2.3
	−190	1.00×10^6	27	10

注：石英的测量平行于光轴。

（2）缺陷对声子的散射

在限制平均自由程方面，几何效应也是很重要的，必须考虑由晶体边界、杂质、缺陷等引起的散射。

点缺陷对声子的散射，取决于点缺陷相对于声子波长来说有多大。点缺陷的尺寸越小，散射越小，平均自由程越长。

在低温下，前面描述过的倒逆过程和点缺陷等杂质散射都很小。所以平均自由程将增加，如果平均自由程 l 可以与样品的宽度相比拟时，则 l 值将受到这个宽度的限制，它不再是温度的函数。于是声子平均自由程为常值。在式（4-78）中，热导率唯一依赖于温度的因子是质量定容热容 c_V，在低温下 c_V 按 T^3 变化，所以 λ

$\propto T^3$。图 4-21 给出了低温下氟化锂晶体的热导率 λ。从图中可以看出，热导率受到表面散射的限制。在 $T < 10\text{K}$ 时，热导率 $\lambda \propto T^3$，来自于低温质量定容热容 c_V 对温度的依赖关系，而且样品尺寸越大，热导率也越大。当温度再升高时，倒逆过程开始起作用。热导率随温度升高而迅速减小，开始以 $e^{\Theta_D/2T}$ 形式，之后以 $1/T$ 形式减少。注意在声子之间散射的平均自由程与样品尺寸相当时，热导率达最大值。

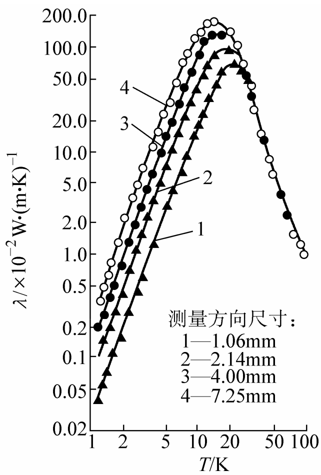


图 4-21　氟化锂晶体的热导率

2. 光子热导

固体中除了声子的热传导外，还有光子的热传导。这是因为固体中分子、原子和电子的振动、转动等运动状态的改变，会辐射出频率较高的电磁波。这类电磁波覆盖了一较宽的频谱。其中具有较强热效应的是波长在 0.4～40μm 间的可见光与部分近红外光的区域。这部分辐射线就称为热射线。热射线的传递过程称为热辐射。由于它们都在光频范围内，其传播过程和光在介质（透明材料、气体介质）中传播的现象类似，也有光的散射、衍射、吸收和反射、折射。所以可以把它们的导热过程看做是光子在介质中传播的导热过程。

在温度不太高时，固体中电磁辐射能很微弱，但在高温时就明显了。因为其辐射能量与温度的四次方成正比。例如，在温度 T 时黑体单位容积的辐射能 E_T 为

$$E_T = 4\sigma n^3 T^4 / c \qquad (4\text{-}83)$$

式中，σ 是斯忒藩—玻耳兹曼常量（为 $5.67\times10^{-8}\ \text{W/(m}^2\cdot\text{K}^4))$；$n$ 是折射率；c 是光速（$3\times10^8\text{m/s}$）。

由于辐射传热中，质量定容热容相当于提高辐射温度所需的能量，所以

$$c_V = \left(\frac{\partial E}{\partial T}\right) = \frac{16\sigma n^3 T^3}{c} \qquad (4\text{-}84)$$

同时辐射线在介质中的速度 $v_r = c/n$，以及将式（4-84）代入式（4-78），可得到辐射能的热导率 λ_r

$$\lambda_r = \frac{16}{3}\sigma n^2 T^3 l_r \qquad (4\text{-}85)$$

l_r 是辐射线光子的平均自由程。

对于介质中辐射传热过程，可以定性地解释为：任何温度下的物体既能辐射出一定频率的射线，同样也能吸收类似的射线。在热稳定状态，介质中任一体积

元平均辐射的能量与平均吸收的能量相等。当介质中存在温度梯度时，相邻体积间温度高的体积元辐射的能量大，吸收的能量小；温度较低的体积元正好相反，吸收的能量大于辐射的，因此，产生能量的转移，整个介质中热量从高温处向低温处传递。λ_r 就是描述介质中这种辐射能的传递能力。它极为关键地取决于辐射能传播过程中光子的平均自由程 l_r。对于辐射线是透明的介质，热阻很小，l_r 较大；对于辐射线不透明的介质，l_r 很小；对于完全不透明的介质，$l_r=0$，在这种介质中，辐射传热可以忽略。一般，单晶和玻璃对于辐射线是比较透明的，因此在 773～1273K 辐射传热已很明显，而大多数烧结陶瓷材料是半透明或透明度很差的，其 l_r 要比单晶和玻璃的小得多，因此，一些耐火氧化物在 1773K 高温下辐射传热才明显。

光子的平均自由程除与介质的透明度有关外，对于频率在可见光和近红外光的光子，其吸收和散射也很重要。例如，吸收系数小的透明材料，当温度为几百摄氏度时，光辐射是主要的；吸收系数大的不透明材料，即使在高温时光子传导也不重要。在无机材料中，主要是光子的散射问题，这使得 l_r 比玻璃和单晶都小，只是在 1500℃以上，光子传导才是主要的，因为高温下的陶瓷呈半透明的亮红色。

3. 电子热导

人们在研究纯金属的热导率时，发现一个引人注意的事实，就是金属的电导率愈高，热导率也就愈高。在室温时，金属的热导率和金属的电导率的比值是一常数，不随金属不同而改变。这就是维德曼（Wiedeman）—弗兰兹（Franz）定律。洛伦兹（Lorentz）研究不同的温度 T 下金属热导率 λ 和金属的电导率 σ 的关系时，发现在各温度下 λ 与 σ 的比值被相应的绝对温度 T 除以后，得到的数值对各金属和各温度都是常数。这个常数就叫做洛伦兹常数，以 L 表示。即

$$\frac{\lambda}{\sigma T}=L \tag{4-86}$$

从经典自由电子论可以推导出洛伦兹关系，简略地叙述如下。

经典电子论对电导率的表达式为

$$\sigma=\frac{1}{2}\frac{e^2 n_0 l}{m\bar{u}} \tag{4-87}$$

式中，n_0 是单位体积的传导电子数；e 是电子的电荷；l 是电子的平均自由程；$\bar{u}$ 是传导电子的平均速度；m 是电子的质量。

单位体积物质的电子定容热容为 $c_V^e=(3/2)n_0k$，这里 k 为玻耳兹曼常数。将此式代入式（4-78）得电子的热导率 λ_e 为

$$\lambda_e=\frac{1}{2}n_0kl\bar{u} \tag{4-88}$$

由式（4-87）和式（4-88）得

$$\frac{\lambda_e}{\sigma}=\frac{\frac{1}{2}n_0kl\bar{u}}{\frac{1}{2}\frac{e^2n_0l}{m\bar{u}}}=\frac{m\bar{u}^2}{e^2}k \tag{4-89}$$

而，$(1/2)m\bar{u}^2=(3/2)kT$，所以可得

$$\frac{\lambda_e}{\sigma T}=\frac{3k}{e^2}=L \tag{4-90}$$

由此可知，当温度一定时，各金属的热导率对电导率之比等于一个相同的常数，而与金属种类无关。在国际单位制中 $L=2.45\times10^{-8}\ \mathrm{W\cdot\Omega/K^2}$。很多实验工作证明：洛伦兹常数只在较高的温度（>0℃）时才近似为常数，在趋向 0K 时，洛伦兹常数也趋近于零，这是由于金属中热的传递不仅借助自由电子实现，而且也有赖于离子的振动。另外，对于大多数金属 L 是符合的，但也不是符合得非常好，不同金属的 L 有些差别，对于合金差别更大。

如果从总的热导率 λ 中减去声子热导率 λ_t，这样式（4-90）就变为

$$\frac{\lambda-\lambda_t}{\sigma T}=L' \tag{4-91}$$

式中 L'是修正后的洛伦兹常数，对所有金属都是常数。修正后的洛伦兹常数对绝大多数金属符合得很好，但只对 Be、Cu 例外。它对合金也可适用。

即便上述两个规律是近似的，但建立了电导率和热导率之间的联系，也是很有意义的，因为测定热导率较测定电导率要困难，而且结果的准确度也较差。另外，电导率的变化解释起来比热导率较简单。

4.6.3　影响热导率的因素

由于在材料中热传导机构和过程是很复杂的，因而对于热导率的定量分析十分困难，下面对影响热导率的一些主要因素进行定性的讨论。

1. 温度的影响

在温度不太高的范围内，主要是声子传导，热导率由式（4-78）给出。其中 $\bar{v}$ 通常可以看成常数，只有在温度较高时，由于介质的结构松驰而蠕变，使介质的弹性模量迅速下降，$\bar{v}$ 减小，如一些多晶氧化物在温度高于 973～1273K 时就出现这一效应。热容 c_V 在低温下与 T^3 成比例，在超过德拜温度便趋于一恒定值。声子平均自由程 l 随着温度升高而降低。实验指出，l 值随温度的变化规律是：低温下 l 值的上限为晶粒的线度；高温下 l 值的下限为晶格间距。

图 4-22 是铁的热导率与温度的关系曲线。在很低温度下声子的平均自由程 l 增大到晶粒的大小，达到了上限，因此，l 值基本上无多大变化。热容 c_V 在低温下与温度的三次方成正比，因此 λ 也近似与 T^3 成比例地变化。随着温度的升高，

λ 迅速增大，然而温度继续升高，l 值要减小，c_V 随温度 T 的变化也不再与 T^3 成比例，并在德拜温度以后，趋于一恒定值。而 l 值因温度升高而减小，成了主要影响因素。因此，λ 值随温度升高而迅速减小。这样，在某个低温处（≈40K），λ 值出现极大值。在更高的温度，由于 c_V 已基本上无变化，l 值也逐渐趋于下限，所以随温度的变化 λ 又变得缓和了。在达到高温后，λ 值又有少许回升，这就是高温时辐射传热带来的影响。铁磁性金属或合金的热导率与温度的曲线在居里点时有转折。

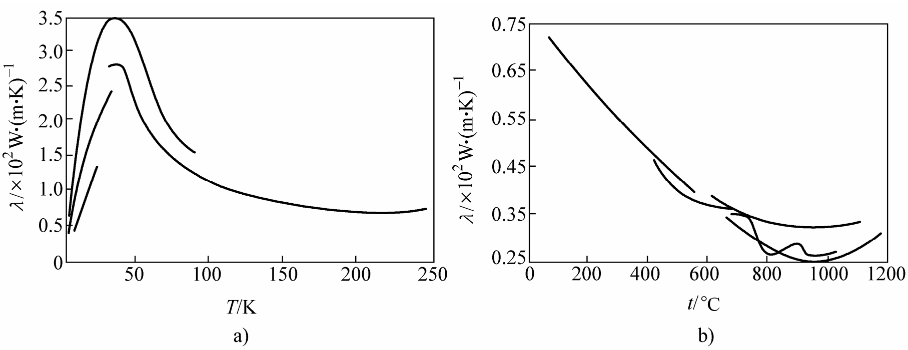


图 4-22　是铁的热导率与温度的关系曲线（不同曲线为不同研究者的曲线）

a）低温下　b）较高温度下

物质种类不同，热导率随温度变化的规律也有很大不同。例如，各种气体随温度上升热导率增大；金属材料在温度超过一定值后，热导率随温度的上升而缓慢下降；耐火氧化物多晶材料在实用的温度范围内，随温度的上升，热导率下降；至于不密实的耐火材料，如粘土砖、硅藻土砖、红砖等，气孔导热占一定分量，随着温度的上升，热导率略有增大；非晶体材料的热导率随温度变化的曲线是另一种情况。

非晶体的导热机理和规律与晶体的有所不同，以玻璃作为一个实例来进行分析。玻璃具有近程有序、远程无序的结构。在讨论它的导热机理时，近似地把它当做由直径为几个晶格间距的极细晶粒组成的“晶体”。这样，就可以用声子导热的机理来描述玻璃的导热行为和规律。从前面晶体的声子导热的机理中，已知声子的平均自由程由低温下的晶粒直径大小变化到高温下的几个晶格间距的大小。因此，对于上述晶粒极细的玻璃来说，它的声子平均自由程在不同温度将基本上是常数，其值近似等于几个晶格间距。根据声子导热式（4-78）可知，在较高温度下玻璃的导热主要由热容与温度的关系决定，在较高温度以上则需考虑光子导热的贡献。在中低温（400～600K）以下，光子导热的贡献可忽略不计。声子导热随温度的变化由热容随温度变化的规律决定，即随着温度的升高，热容增大，

玻璃的热导率也相应地上升。这相当于图 4-23 中的 O*F* 段。从中温到较高温度（600～900K），随着温度的不断升高，声子热容不再增大，逐渐为一常数，因此，声子导热也不再随温度升高而增大，因而玻璃的热导率曲线出现一条与横坐标接近平行的直线，这相当于图中的 *Fg* 段。如果考虑此时光子导热在总的导热中的贡献已开始增大，则为图 4-23 中的 *Fg′*段。高温以上（超过 900K），随着温度的进一步升高，声子导热变化仍不大，这相当于图 4-23 中的 *gh* 段。但由于光子的平均自由程明显增大，根据式（4-85），光子热导率将随温度的三次方增大。此时光子热导率曲线由玻璃的吸收系数、折射率以及气孔率等因素决定。这相当于图 4-23 中的 *g′h′*段。对于那些不透明的非晶体材料，由于它的光子导热很小，不会出现 *Fh′*这一段线。

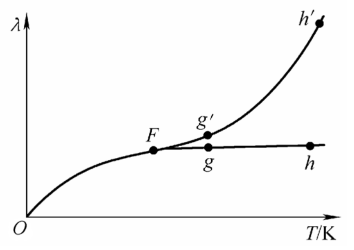


图 4-23　非晶体热导率曲线

总的来说，非晶体的热导率（不考虑光子导热的贡献）在所有温度下都比晶体的小；两者在在高温时比较接近；非晶体的热导率没有峰值点。

2. 显微结构的影响

（1）结晶构造的影响

声子传导与晶格振动的非谐性有关。晶体结构愈复杂，晶格振动的非谐性程度愈大，格波受到的散射愈大，因此，声子平均自由程较小，热导率较低。

（2）单质多晶体与单晶体的热导率

对于同一种物质，多晶体的热导率总是比单晶小。这是由于多晶体中晶粒尺寸小，晶界多，缺陷多，晶界处杂质也多，声子更易受到散射，它的平均自由程小得多，所以热导率小。低温时多晶体的热导率与单晶的平均热导率一致，但随着温度升高，差异迅速变大。这也说明了晶界、缺陷、杂质等在较高温度下对声子传导有更大的阻碍作用，同时也是由于单晶体在温度升高后比多晶体在光子传导方面有更明显的效应。

3. 化学组成的影响

不同组成的晶体，热导率往往有很大差异。这是因为构成晶体的质点的大小、性质不同，它们的晶格振动状态不同，传导热量的能力也就不同。一般说来，质点的相对原子质量愈小，密度愈小，弹性模量愈大，德拜温度愈高，则热导率愈大。这样，轻元素的固体和结合能大的固体热导率较大，如金刚石的 $\lambda=1.7\times10^{-2}$W/(m·K)，较轻的硅、锗的热导率分别为 $\lambda=1.0\times10^{-2}$W/(m·K)和 $\lambda=0.5\times10^{-2}$W/(m·K)。

晶体中存在的各种缺陷和杂质会导致声子的散射，降低声子的平均自由程，使热导率变小。固溶体的形成同样也降低热导率，而且取代元素的质量和大小与基质元素相差愈大，取代后结合力改变愈大，则对热导率的影响愈大。这种影响

在低温时随着温度的升高而加剧。当温度高于德拜温度的一半时，与温度无关。这是因为极低温度下，声子传导的平均波长远大于线缺陷的线度，所以并不引起散射。随着温度升高，平均波长减小，在接近点缺陷线度后散射达到最大值，此后温度再升高，散射效应也不变化，从而与温度无关了。图 4-24 表示了 MgO-NiO 固溶体热导率与组成的关系。在杂质含量很低时，杂质效应十分显著。所以在接近纯 MgO 或纯 NiO 处，杂质含量稍有增加，λ 值迅速下降。随着杂质含量的增加，这个效应不断减弱。另外，从图中还可以看到，杂质效应在 200℃比 1000℃要强。若低于室温，杂质效应会更强烈。

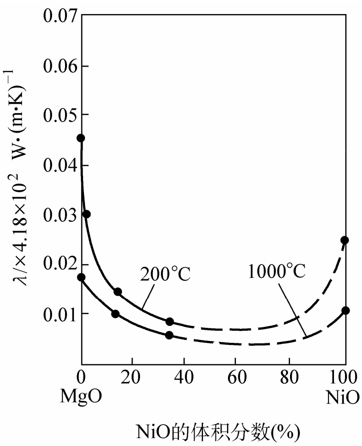


图 4-24　MgO-NiO 固溶体热导率

4. 复相材料的热导率

常见的材料典型微观结构是分散相均匀地分散在连续相中，例如，陶瓷晶相分散在连续的玻璃相中。这个类型的材料的热导率 λ 可按下式计算

$$\lambda = \lambda_c \frac{1+2V_d\left(1-\lambda_c/\lambda_d\right)/\left(2\lambda_c/\lambda_d+1\right)}{1-V_d\left(1-\lambda_c/\lambda_d\right)/\left(2\lambda_c/\lambda_d+1\right)} \tag{4-92}$$

式中，λ_c、λ_d 分别为连续相和分散相物质的热导率；V_d 为分散相的体积分数。

无机材料常含有气孔，气孔对热导率的影响较为复杂。一般当温度不很高，而且气孔率不大，气孔尺寸很小，又均匀地分散在陶瓷介质中时，这样的气孔可看做为一分散相，陶瓷材料的热导率仍然可以按式（4-92）计算。只是，因为与固体相比，它的热导率很小，可近似看做为零。Eucken 根据式（4-92），因气相 $\lambda_d \approx 0$，固相为连续相，$Q=\lambda_c/\lambda_d$ 很大，则该式成为

$$\lambda = \lambda_c \frac{1+2V_d\left(1-Q\right)/\left(2Q+1\right)}{1-V_d\left(1-Q\right)/\left(2Q+1\right)} = \lambda_c \frac{2Q\left(1-V_d\right)}{2Q\left[1+(1/2)V_d\right]} \approx \lambda_c\left(1-V_d\right) = \lambda_s(1-p) \tag{4-93}$$

式中，λ_s 是固相的热导率；p 是气孔的体积分数。更精确一些的计算是在式（4-93）的基础上，再考虑气孔的辐射传热，导出公式为

$$\lambda = \lambda_s(1-p) + \frac{p}{(1-p_L)/\lambda_s + p_L/(4G\varepsilon\sigma dT^3)} \tag{4-94}$$

式中，p 是气孔的面积分数；p_L 是气孔的长度分数；ε 是辐射面的热发射率；G 是几何因子。顺向的长条气孔，$G=1$；横向的圆柱形气孔，$G=\pi/4$；球形气孔，$G=$

2/3。d 是气孔的最大尺寸。当热发射率 ε 较小，或温度低于 500℃时，可直接使用式（4-93）。

含有微小气孔的多晶陶瓷，其光子自由程显著减小，因此，大多数无机材料的光子传导率要比单晶和玻璃的小 1～3 个数量级。光子热传导效应只有在温度大于 1773K 时才是重要的；另一方面，少量的大气孔对热导率影响较小，而且当气孔尺寸增大时，气孔内气体会因对流而加强传热。当温度升高时，热辐射的作用增强，它与气孔的大小和温度的三次方成比例。这一效应在温度较高时，随温度的升高加剧。这样气孔对热导率的贡献就不可忽略，式（4-93）也就不适用了。

粉末和纤维材料的热导率比烧结材料的低得多。这是因为在其间气孔形成了连续相。材料的热导率在很大程度上受气孔相热导率所影响。这也是粉末、多孔和纤维类材料有良好热绝缘性能的原因。

第 5 章　材料的电导性能

当今世界，电能的开发与电信的发展已深入到社会生产和家庭生活的每一个角落，影响着我们的整个生产和生活方式，“电气化”曾被作为一个社会物质文明的重要指标。从大功率发电机、变压器、几千公里的电能输送到微电子线路的各种元件，无一不在应用着材料的不同电学性能。它们分别被制成导体材料，电阻和电热材料，热电、压电和光电材料，半导体材料和绝缘材料等等。特别是作为 20 世纪的十大发明之一，半导体材料的发展导致了大规模集成电路的出现，推动了电子计算机技术的巨大进步，使人类社会的生产和生活发生了深刻的变化。另外，超导电性的发现和超导材料的研究与进展，又对强磁场的获得和无铜耗电机等技术提供了新的发展途径。不言而喻，研究材料的电学性能是多么的重要。

材料的电学性能大致上包含有导电性、超导电性、介电性、磁电性、热电性、接触电性、热释电性和压电性、光电性等等。在材料的许多应用中，电导性能是十分重要的。

5.1　金属材料的电导性能

一般说金属具有良好的导电性能。金属材料的电学性能依其成分、原子结构、能带结构、组织状态而异。外界因素（诸如温度、压力、形变、热处理等）通过改变金属材料内部结构或组织状态而影响其电导性能。本节将介绍金属材料电导性能的一些重要规律。

5.1.1　金属材料电导机制与马基申定理

由量子自由电子理论和能带理论可知材料的电导率 σ 为

$$\sigma = \frac{n^* e^2}{2m^*}\bar{\tau} \tag{5-1}$$

式中，e 为电子的电荷量；m^*为电子的有效质量（它是考虑晶体点阵周期场势与电子相互作用的结果）；n^*为单位体积内实际参加导电的电子数（有效电子数）（即能量接近费米面的那部分电子而非单位体积内全部的电子数）；$\bar{\tau} = \bar{L}/\bar{\upsilon}$ 为相邻两次散射的时间间隔，$\bar{L}$ 为电子的平均自由程，$\bar{\upsilon}$ 为电子运动的平均速度。式（5-1）适用于所有材料，它能完整地反映晶体导电的物理本质。量子力学可以证明，当电子波在绝对零度下通过一个理想的晶体点阵时，它将不会受到散射而无阻碍地

传播，这时电阻率 $\rho = 0$（$\rho = 1/\sigma$），而 σ 为无穷大，即此时的材料是一个理想的导体。只有在晶体点阵的完整性以及由于晶体点阵离子的热振动，晶体中的异类原子、位错和点缺陷等使晶体点阵的周期性遭到破坏的地方，电子波才会受到散射，从而产生了阻碍作用，降低了导电性，这就是材料产生电阻的本质所在。

传导电子的散射发生在电子—声子、电子—杂质原子以及电子与其他晶体点阵静态缺陷相碰撞的时候。理想金属的电阻对应着两种散射机制（声子散射和电子散射，在极低温下“电子—电子”散射对电阻的贡献可能是显著的，但除低温以外几乎所有温度下大多数金属的电阻都取决于“电子—声子”散射。），可以看成为基本电阻。这个电阻在绝对零度时降为零。第三种机制（电子在杂质和缺陷上的散射）在有缺陷的晶体中，这种机制形成的电阻即为绝对零度下金属的残余电阻。这个电阻反映了金属的纯度和完整性。

马基申等人指出，若固溶体中溶质原子的浓度较小时，则可把固溶体电阻率看成由金属基本电阻率 $\rho(T)$和残余电阻率 ρ_r 组成（马基申定律），即

$$\rho = \rho(T) + \rho_r \tag{5-2}$$

式中，$\rho(T)$为与温度 T 有关的金属的基本电阻率，即溶剂金属（纯金属）的电阻率；ρ_r 为决定于化学缺陷和物理缺陷而与温度无关的残余电阻率。这里所指的化学缺陷是偶然存在的杂质原子及人工加入的合金元素的原子，物理缺陷是指空位、间隙原子、位错以及它们的复合体。

5.1.2　影响纯金属导电的因素

1. 温度对纯金属电阻的影响

在无组织变化的条件下，温度对金属电阻的影响是由于温度引起离子晶格热振动造成对电子波的散射，而使电阻率随温度的升高而增加。当晶体为理想完整时，在绝对零度下，因为没有温度引起的离子晶格热振动所造成的电子波散射，故电阻率为零。大量的实验证明，许多纯金属的电阻率在很宽的温度范围内，可用如下经验公式描述

$$\rho(T) = \frac{AT^5}{M{\Theta_D}^6}\int_0^{\Theta_D/T}\frac{x^5\mathrm{d}x}{(\mathrm{e}^x - 1)(1-\mathrm{e}^{-x})} \tag{5-3}$$

式中，A 为金属的特性常数；M 为金属原子质量；$x = h\nu/(kT)$为积分变数，其中 h 为普朗克常数，ν 为原子热振动频率，k 为玻尔兹曼常数；Θ_D 为德拜温度，$\Theta_D = h\nu_m/k$，其中 ν_m 为原子热振动频率的最大值。Θ_D 为德拜温度，不同元素具有不同的德拜特性温度，如 Al、Mn 元素的德拜特性温度分别为 428K 和 450K。

当 $T > 0.5\Theta_D$ 时，上式简化成

$$\rho(T) = \frac{A}{4}\frac{T}{M{\Theta_D}^2} \tag{5-4}$$

这就是熟知的金属在高温下的电阻率同温度成正比的关系。

在很低的温度下，即 $T < \Theta_D$，则式（5-3）的积分上限可以认为 $\Theta_D/T \to \infty$，这样所得的积分值为 124.4，此时金属的电阻率

$$\rho(T) = 124.4\frac{AT^5}{M{\Theta_D}^6} \tag{5-5}$$

即电阻率同温度的五次方成正比。

可以从电子受格波（声子）散射机制来解释上述经验规律。这里我们简单地解释高温下的电阻率同温度成正比的关系。当晶体的温度升高，晶格离子的热振动增强，这将增加离子对电子的散射几率，因而电阻率随温度上升而增大。下面用简单的模型讨论晶格的振动对电阻的影响。如果晶格的振动频率 v 与爱因斯坦温度 Θ_E 之间有温度关系 $hv = k\Theta_E$，此处 k 为玻耳兹曼常数。设原子的质量为 M，离开平衡位置的距离为 x，恢复力常数为 β，其运动方程为

$$M\frac{\mathrm{d}^2x}{\mathrm{d}t^2} = \beta x \tag{5-6}$$

那么角频率 ω 为

$$\omega^2 = \frac{\beta}{M} = 4\pi^2v^2 = 4\pi^2\left(\frac{k}{h}\Theta_E\right)^2 \tag{5-7}$$

原子的振动势能 $U = (1/2)\beta x^2$，设其振动振幅均方为 A^2，其平均势能为

$$\bar{U} = \frac{1}{2}\beta A^2 = \frac{\beta}{2}\frac{\int_{-\infty}^{\infty} x^2 \mathrm{e}^{-v/(kT)}\mathrm{d}x}{\int_{-\infty}^{\infty} \mathrm{e}^{-v/(kT)}\mathrm{d}x} = \frac{1}{2}kT \tag{5-8}$$

则

$$A^2 = \frac{kT}{\beta} = \frac{h^2}{4\pi^2 Mk\Theta_E^2}T \tag{5-9}$$

A^2 愈大，表示振动愈强烈，电子散射愈强，散射几率 P 愈大，故可以认 $P \propto A^2$，$\rho \propto T$。这说明了高温时电阻率与绝对温度成正比。碱金属与贵金属是一价金属，$\sigma/(M{\Theta_E}^2)$的值最大，二价金属的较低，过渡金属的值为最低。二价金属其值较低的原因是 s 带与较高的导带有小部分重迭，金属性主要来自能带的重迭，对导电有作用的粒子，是导带中的电子及满带（s 带）的空穴，它们的浓度都比较小，故电导率较低。而过渡金属，对导电有贡献的粒子是 4s 带的 s 电子及 3d 带中的空穴，其电导率应为

$$\sigma = \frac{n_e e^2}{2{m_e}^*}\bar{\tau}_e + \frac{n_h e^2}{2{m_h}^*}\bar{\tau}_h \propto \frac{s_e}{{d_h}^{1/3}} \tag{5-10}$$

式中，n_e、n_h分别为 s 带参加导电的电子数及 d 带中参加导电的空穴数；m_e^*、m_h^*分别为 s 带参加导电的电子的有效质量及 d 带中参加导电的空穴的有效质量；s_e、d_h分别为 s 带的电子数及 d 带中的空穴数；$\overline{\tau}_e$、$\overline{\tau}_h$分别为电子和空穴的平均自由时间。由于 3d 带比较窄，能级密度大，散射几率大，m^*也大，故 σ 小。

若以 ρ_t 和 ρ_0 分别表示材料在 0℃和 t℃下的电阻率，则电阻率与温度的关系可表示为一个幂函数

$$\rho_t = \rho_0(1+\alpha t+\beta t^2+\gamma t^3+\cdots) \tag{5-11}$$

这里，α、β、γ 为电阻温度系数。实验证明，对于普遍非过渡族金属 Θ_D 一般不超过 500K，当温度超过 $2/3\Theta_D$ 时，β、γ 及高次项系数很小，线性关系已有足够的精度，即在室温和更高一些温度可写成

$$\rho_t = \rho_0(1+\alpha t) \tag{5-12}$$

除过渡族金属外，所有纯金属的电阻温度系数 α 都近似等于 4×10^{-3}。过渡族金属特别是铁磁性金属具有较高的 α 值，如铁的 α=6×10^{-3}。

多晶型金属不同的晶型导致了同一金属存在不同的物理性能，其中包括电阻率与温度的关系。由于不同晶型的电阻温度系数明显不同，在 $\rho-t$ 曲线上多晶转变可以显示出来，出现明显的反常。对于铁磁性金属及其合金，电阻和温度的线性关系已不适用。在居里点以下 ρ 与 t 的关系偏离线性更为显著。铁磁性金属在接近居里点的电阻率反常量 $\Delta\rho$ 与自发磁化强度 I 的平方成正比，即 $\Delta\rho/\rho_0\propto I^2$，此处 ρ_0 为在居里点以下的电阻值。这一反常现象与自发磁化中 s 电子与 d 电子的相互作用有关。

2. 纯金属中的缺陷对导电性的影响

金属中的各种缺陷造成晶格畸变引起电子波散射，从而影响导电性。位错与点缺陷（空位及间隙原子）相比对 ρ 的贡献极小。所以在研究缺陷对 ρ 的影响时，主要应研究点缺陷的影响。金属中空位的浓度大小主要由温度的高低决定。金属在任何温度下，总存在着线缺陷（位错）与点缺陷的平衡浓度。在任何温度下，空位的形成能均较其他缺陷的低，故空位的浓度高，它对 ρ 的影响也最大。金属中空位的浓度 C_h 与热力学温度 T 的关系可用下式描述

$$C_h = C_0\exp(-E_h/kT) \tag{5-13}$$

式中，E_h 为形成一个空位的能量；C_0 为常数；k 为玻耳兹曼常数。影响 C_h 的一个因素是温度，另一个因素是原子结合力的强弱（影响 E_h 的大小）。例如，在室温下难熔金属的 C_h 就比中等熔点金属的 C_h 低得多。造成金属中缺陷的原因有多种多样，如幅照、冷热加工、热处理以及各种工艺过程及使用过程等都可能造成金属中的缺陷。

塑性变形过程中形成点缺陷与位错，因而 ρ 增大，其增大数值 $\Delta\rho$ 与变形量 ε 有关，如下式所示

$$\Delta\rho = C\varepsilon^n \tag{5-14}$$

式中，C 为比例系数；n 在 0～2 范围内。纯金属经大变形量冷加工后（如 Al、Cu、Fe 等），在室温下电阻率 ρ 增大仅为 2%～6%。W 是个例外，经冷加工后 ρ 可增大百分之几十。ρ 增大的原因首先是由于晶格畸变所致，同时冷加工也改变原子结合力并可导致原子间距增大。经冷加工后的金属再进行退火，则 ρ 下降，若退火温度高于再结晶温度，则 ρ 可恢复到初始值，这是因为在回复及再结晶过程中，冷加工所造成的晶格畸变及各种缺陷逐渐消除之故，在退火时若发生相变，ρ 将发生显著变化。淬火对纯金属的 ρ 有明显的影响，淬火使 ρ 明显增大，因为淬火过程中引入了大量空位。

3. 受力状态对金属电阻的影响

在弹性范围内，单向拉伸或扭转应力能提高金属的电阻率 ρ，并有

$$\rho = \rho_0(1 + \alpha_\gamma\sigma_t) \tag{5-15}$$

式中，ρ_0 是无负荷时金属的电阻率；α_γ 是拉应力系数；σ_t 是拉应力，单位为 Pa。

在弹性范围内，对大多数金属来说，在受压应力情况下电阻率 ρ 降低，有

$$\rho = \rho_0(1 + \varphi\sigma_p) \tag{5-16}$$

式中，ρ_0 是真空下的电阻率；φ 是压应力系数，为负值；σ_p 是压应力，单位为 Pa。

压力之所以能对材料的性能表现出如此强烈的影响，是由于金属在压力的作用下，其原子间距缩小，内部缺陷的形态、电子结构、费米面和能带结构以及电子散射机制等都将发生变化，这必然会影响金属的导电性能。尤其对过渡族金属，由于其内部存在着具有能量差别不大的未填满电子的壳层，因此在压力的作用下，有可能使外壳层电子转移到未填满的内壳层，这就必然会表现出性能的变化。

在不同温度下，人们几乎对所有纯元素都研究过压力对电阻的影响，并确定了电阻的压力系数（定义为$(1/\rho)(\mathrm{d}\rho/\mathrm{d}\sigma_p)$）。实验表明，随温度的变化电阻压力系数几乎不变，这说明电阻压力系数与温度无关。根据压力对电阻的影响，可把金属分为两类，一类是正常金属元素，其电阻率随压力增大而下降，如 Fe、Co、Ni、Rh、Pd、Ir、Pt、Cu、Ag、Zr、Hf、Th、Nb、Ta、M、W、V 等。另一类为反常金属元素，它们是碱金属、碱土金属、稀土金属和第 V 族的半金属，它们有正的电阻压力系数，但随压力升高一定值后系数变号。研究表明，这种反常现象和压力作用下的相变有关。

高压力往往能导致物质的金属化，引起导电类型的变化，而且有助于从绝缘体→半导体→金属→超导体的某种转变。

5.1.3　固溶体合金的导电特性

在合金固溶体中，合金元素使晶格发生畸变；改变了能带结构，使费米能位移，改变状态密度及电子有效质量；改变弹性模量，从而影响晶格离子的振动谱。这些都将引起电阻及其他性能变化。不仅如此，固溶体有时会发生同素异构转变、有序无序转变、磁性转变等，这些变化对电阻也有很大影响。

1. 连续固溶体合金的导电特性

由非过渡族金属组成的连续固溶体，对两组元固溶体，A 组元的摩尔浓度为 x、B 组元的摩尔浓度为 $1-x$，则合金的电阻率 ρ 大体与 $x(1-x)$成正比。其电阻最大值通常在 50%摩尔浓度处。而电阻温度系数随浓度的变化刚好与 ρ 相反，在 50%摩尔浓度处有一最小值。这一现象是由于异类原子的存在造成晶格畸变，因而增加了对电子的散射作用。但对固溶体中含有过渡族元素时，ρ 最大值不在 50%摩尔浓度处，而偏向过渡族组元方向。过渡族金属组成固溶体后，其电阻值显著提高（有时增大几十倍）。这是由于过渡金属有未满的 d 或 f 电子壳层，组成固溶体时使得一部分价电子进入未满的 d 或 f 壳层中，使 n_e 减少，故 ρ 增大。这一情况有重要的实用意义，因为目前生产的电热合金和精密电阻合金，绝大部分含有一个以上过渡族元素。在电阻材料中应用最多的过渡族元素是 Fe、Ni、Mn、Cr 等。

固溶体加热时它的电阻通常要增大，但它的电阻温度系数与纯金属相比总是较低的，而且与电阻率一样随成分而变。由非过渡族金属组成的低浓度连续固溶体，由马基申定律可以看出，对同一溶剂的一系列固溶体电阻率与温度曲线的斜率相同，且在常温以上斜率为常数而不取决于杂质的浓度。又从电阻温度系数的定义 $\alpha_T = (1/\rho_T)(\mathrm{d}\rho/\mathrm{d}T)$可知，$\rho_T$ 越大（即存在杂质引起的残余电阻率 ρ_r 时），α_T 就越小。工业上使用的电阻合金的温度系数普遍比导体合金小 1~2 个数量级，也证实了这一规律性。

在低浓度下，加入第二组元所造成的残余电阻率 ρ_r 可表示为 $\rho_r= C\zeta$，这里 ζ 为单位浓度的残留电阻（加入 1%原子异类元素所造成的残留电阻），C 为组元的原子浓度。以一价元素（Cu、Ag、Au）为基的固溶体，其 ζ 随溶质元素（非过渡族元素）的价数增加而增大，可用下式表述

$$\zeta = a + b(z - z_p)^2 \qquad (5\text{-}17)$$

式中，a、b 为常数；z 与 z_p 为溶质及溶剂元素的价数。Mott 的集体化电子理论导出的结果是，固溶体的残留电阻 ρ_r 正比于二组元原子价差的平方，反比于溶剂金属的原子体积。

2. 有序固溶体合金的导电特性

在若干合金体系的固溶体中存在有序-无序转变。固溶体的有序化对合金的电

阻率有显著的影响。如成分相当于 Cu_3Au 和 CuAu 的合金经高温淬火及退火处理后，其电阻值发生了很大的变化。退火后的电阻只有原淬火态的 1/2 或 1/3，如图 5-1 和图 5-2 所示。图 5-2 中曲线 2 上的点离虚线的距离表示有序合金 Cu_3Au 和 CuAu 的残余电阻。

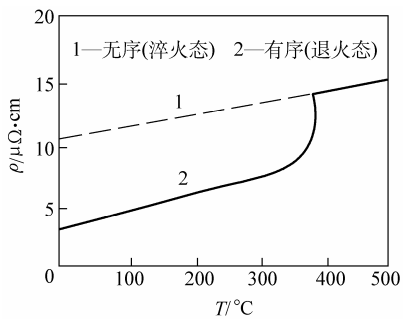


图 5-1　Cu_3Au 合金有序化对电阻率的影响

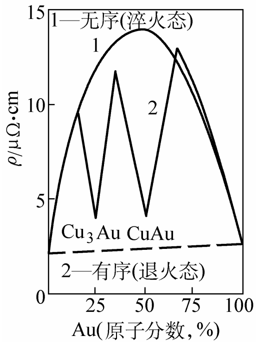


图 5-2　CuAu 合金电阻率曲线

X 射线结构分析发现，对于退火态的 Cu_3Au 和 CuAu 合金，除了代表具有面心立方点阵的无序固溶体的 X 射线谱外，还出现了一些新的线谱，称为超结构线谱。这种新线谱的出现是由于 Cu_3Au 和 CuAu 合金在退火时晶体点阵中的原子进行了有序排列（称为合金固溶体有序化）的缘故。当固容体有序化后，一方面，异类原子使点阵的周期场遭到破坏而使电阻率增大，而固溶体的有序化则有利于改善离子电场的规整性，从而减少电子的散射。另一方面，有序化使组元之间化学作用加强，导致传导电子数目减少。在上述两种相反作用的影响下，电场对称性增加使电阻率下降起着主导作用，所以有序化表现是电阻率降低。

斯米尔诺夫根据合金的成分及远程有序度，从理论上计算了有序合金的残余电阻率。其假定：完全有序合金在 0K 和纯金属一样不具有电阻，只有当原子的有序排列遭到破坏时才有电阻。这样，有序合金的残余电阻率就可写成

$$\rho_r = A\left[c(1-c) - \frac{v}{1-v}(q-c)^2\eta^2\right] \tag{5-18}$$

式中，ρ_r 为 0K 时合金的电阻率；A 为与组元性质有关的系数；c 为合金中第一组元的相对原子浓度；v 为第一类结点（第一组元占据）的相对浓度；对于 $c < v$，$q = c/v$，而对于 $c \geqslant v$，$q = 1$，q 值表示第一类结点被相应原子占据的可能性，即被第一组元原子占据的可能性；η 表示远程有序的程度，$\eta = (p-c)/(q-c)$；p 为被自己原子即第一组元原子占据的第一类结点的相对数目。图 5-3 给出了残余电阻与远程有序度 v 及成分的关系。

有序化不仅存在于以组元为基的一级固溶体中，具有与组元不同空间点阵的

中间相也可能存在有序化。

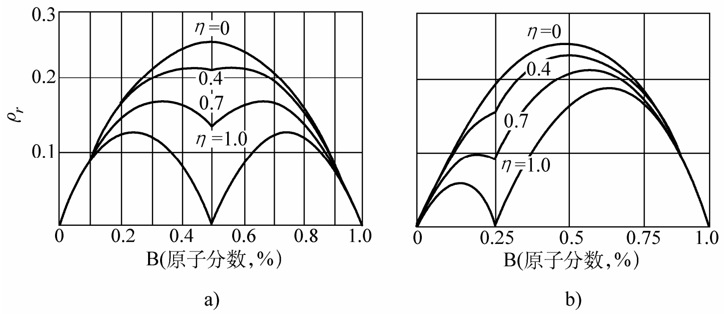


图 5-3 不同程度远程有序（η）下的残余电阻

a）形成 AB 型超结构　b）形成 A_3B 型超结构

3. 不均匀固溶体合金的导电特性

某些含有过渡族金属的合金，如 Ni-Cu、Ni-Cr、Ni-Cu-Zn、Fe-Al、Cu-Mn、Ag-Mn、Au-Cr 等合金，合金虽是单相固溶体结构，但由于存在特殊相变及特殊结构，其电阻出现反常变化：电阻-温度曲线呈 S 形或部分 S 形；冷加工可使合金的电阻率降低，而退火使电阻率升高；高温淬火处理后再经低温回火，室温电阻上升。把这种电阻率反常增大效应称为“K 状态”。在“K 状态”形成时，合金虽然处于单相固溶体状态，但固溶体内组元原子在晶体中分布不均匀，由 X 射线漫散射已证实其原子间距的大小显著波动。固溶体内形成了原子的偏集区，偏集区范围大约有 100 个原子，形成不均匀固溶体状态。由于偏聚区造成电子波的附加散射，使固溶体电阻率增大，一般可增加 10%～15%。也有人认为“K 状态”标志着某种短程有序。继续升高温度或冷加工变形将破坏这些原子偏集区，从而驱散了“K 状态”，使电阻率恢复正常态。

5.1.4　金属化合物、中间相及多相合金导电性

1. 金属化合物

若组元间的电负性相差较大，原子间的键合具有离子键的性质，则在许多情况下，均形成金属化合物。金属化合物的电导率比较小，一般情况下，它比形成化合物的组元的电导率要小得多。因为形成化合物后，原子间结合类型发生变化，原子间的金属结合至少部分地变为共价结合，甚至是离子结合，载流的电子浓度减少。有时组成化合物之后，合金变成半导体材料。半导体型的金属化合物在加热时，电阻率下降，其电导率与温度关系按指数规律变化。在金属型导电的化合物中，电阻随温度升高而增大，如 HgSe。金属化合物的电导率与其组元之间电离势之差有关，此差值减小则电导率增大。

2. 中间相

中间相包括电子化合物、间隙相等。电子化合物的电阻值随温度升高而增大，在熔化时电阻值下降。间隙相主要是指过渡族金属与氢、氮、碳、硼组成的化合物。非金属元素处在金属原子点阵的间隙之中，这类相绝大部分是属于金属型的化合物，具有明显的金属导电性，其中一些（例如 TiN、ZrN）是良好的导体，比相应的金属组元的导电性还好。这些相的正电阻温度系数与固溶体电阻温度系数有相同的数量级。这些相具有金属性的结合，并且非金属给出部分价电子到传导电子中去，这是大部分电子化合物导电性好的原因。

3. 多相合金

当合金由两个以上相组成时，其合金的导电性是由组成相的导电性决定的。由于导电性是组织敏感参数，晶粒大小、晶界状态及织构等因素均对导电性产生影响。另外，若一种相的尺寸与电子平均自由程为相同数量级时，则此时对电子产生最大的散射作用（电阻率可升高 10%～15%）。如果这些因素都可以忽略的话，则两相（或多相）合金的电阻可以从各相电阻的算术相加而求得。若合金经充分退火，无组织机构，由比较大的等轴晶组成，且各相的电导率相近（电导率之比为 0.75～1.75）时，合金的电导率与组元（相）的体积浓度成线性关系（由电阻的算术和而得），同时，电阻率应当按双曲线变化，其表达式为

$$\rho = \rho_1^{\,p} \rho_2^{\,q} \tag{5-19}$$

式中，ρ、ρ_1、ρ_2 为该合金及组元的电阻率；p 及 q 为组元的体积浓度（$p + q = 1$）。

在细小的夹杂物或相（10^{-6}～10^{-7}cm）以及沿晶界的点阵畸变所引起的电子附加散射不计入的条件下，若已知合金中每个相的体积浓度、形状及相互分布，原则上可按各相的电导率来计算合金的电导率。也可以用下式计算合金的电导率

$$\sum_i \frac{\sigma_i - \sigma}{\sigma_i - 2\sigma} c_i = 0 \tag{5-20}$$

式中，σ、σ_i 分别为合金及组元 i 的电导率；c_i 为组元 i 的体积分数。对常常遇到的基体型组织（基体中分布着互不相连的第二相），式（5-20）经过某些修正可得到与实验结果符合很好的公式

$$\sigma = \sigma_0 \left(1 - 3nc + 3n^2 c^2 \frac{\sigma_0 + 2\sigma_1}{\sigma_1 + 2\sigma_0} \right) \tag{5-21}$$

式中，$n = (\sigma_0 - \sigma_1)/(2\sigma_0 + \sigma_1)$；$\sigma$ 为合金的电导率；σ_0 和 σ_1 分别为基体与第二相的电导率；c 为第二相的体积分数。对统计型组织（由混乱分布两相所组成的不规则混合体），合金的电导率还可以用下式描述

$$\sigma = \frac{(3c_1 - 1)\sigma_1 + (3c_2 - 1)\sigma_2}{4} + \left\{ \frac{[(3c_1 - 1)\sigma_1 + (3c_2 - 1)\sigma_2]^2}{16} + \frac{\sigma_1 \sigma_2}{2} \right\}^{1/2} \tag{5-22}$$

式中，下角号 1、2 分别表示在其组织（混合体）中属于第一相或第二相。当合金中析出相大小不均，或出现纤维状，层状组织时，就不能套用式（5-20）、式（5-21）和式（5-22）。

5.2　半导体导电性能

半导体材料无论是按电阻率（ρ 为 10^{-3}~$10^{9}\Omega\cdot m$）还是按禁带宽度（禁带宽度 ε_g 为 0.2~3.5eV），其电学性能都介于金属导体（$\rho < 10^{-5}\Omega\cdot m$，$\varepsilon_g = 0$eV）与绝缘体（$\rho > 10^{9}\Omega\cdot m$，$\varepsilon_g > 3.5$eV）之间。这也是其“半导体”命名的由来。

半导体材料分为晶体半导体、非晶半导体及有机半导体。晶体半导体又分为元素半导体、化合物半导体和固溶体半导体。非晶半导体主要有非晶硅、多孔硅及硫属玻璃半导体。有机半导体主要是一些高聚物半导体。非晶半导体及有机半导体的研究尚处起步阶段，这里只讨论晶体半导体。

5.2.1　本征半导体

半导体制备工艺的目的之一是制备尽可能纯的材料（纯度可高达 10^{-10}），然后可控制地引入杂质（称为掺杂）。把纯的半导体称为本征半导体，因为它们的行为仅仅由它固有的性质决定（载流子只有半导体晶格本身提供）。把由于外部作用而改变半导体固有性质的半导体称为非本征（杂质）半导体。半导体器件主要使用的是非本征半导体。为了便于理解，首先讨论本征半导体。

以半导体硅为例进行讨论。硅具有金刚石结构，四共价键对称排列。每个原子的四个价电子都参与形成共价键。根据能带理论可知，在 0K 时所有电子都处于价带，导带和价带之间有 1.1eV 的能隙，半导体的空带中无电子，即无运动的电子。但当温度升高或受光照射时，也就是半导体受到热激发时，共价键中的价电子由于从外界获得了能量，其中部分获得了足够大能量的价电子就可以挣脱束缚，离开原子而成为自由电子。反映在能带图上，就是一部分满带中的价电子获得了大于 $\varepsilon_g = 1.1$eV 的能量，跃迁到空带中去。这时空带中有了一部分能导电的电子，称为导带。而满带中由于部分价电子的迁出出现了空位置，如图 5-4 所示。当一个价电子离开原子后，在共价键上留下一个空位（称空穴），在共有化运动中，相邻的价电子很容易填补到这个空位上，从而又出现了新的空穴，其效果等价于空穴移动了，如图 5-5 所示。在无外电场作用下，自由电子和空穴的运动都是无规则的，平均位移为零，所以并不产生电流。但在外电场的作用下，电子将逆电场方向运动，空穴将顺电场方向运动。从能带图可以看出，自由电子在导带内（导带底附近），空穴在价带内（价带顶附近），在本征激发（常见是热激发）过程中它们是成对出现的。在外电场作用下，自由电子和空穴都能导电，所以它们统

称为载流子。

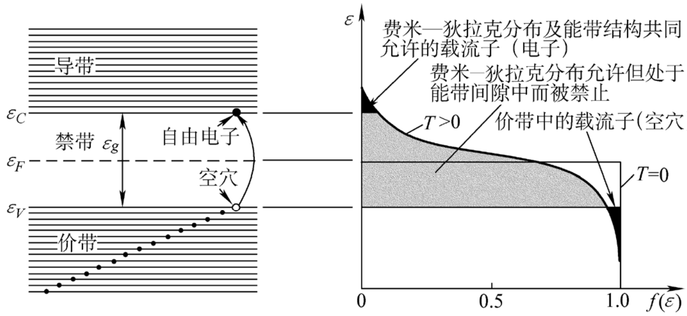


图 5-4　本征激发过程中被占能级及费米—狄拉克分布

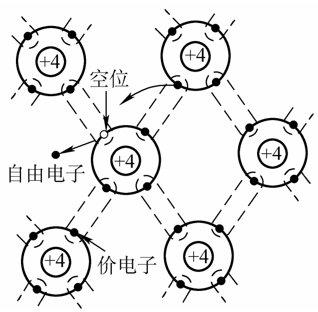


图 5-5　本征激发过程中共价键上的电子运动

1. 本征半导体中载流子数

本征半导体的载流子只有半导体晶格本身提供，根据费米统计，可以计算出导带中电子浓度以及价带中的空穴浓度。

在某一能带（ε_1 和 ε_2 之间），单位体积晶体存在的电子浓度 n_e 可以表示为

$$n_e = \int_{\varepsilon_1}^{\varepsilon_2} \rho(\varepsilon) f(\varepsilon) \mathrm{d}\varepsilon \tag{5-23}$$

式中，$\rho(\varepsilon)$为单位体积晶体能量为 ε 的电子的状态密度；$f(\varepsilon)$为费米—狄拉克分布函数，即能量为 ε 的电子存在的几率。室温下 $kT \approx 0.025\text{eV}$，$\varepsilon - \varepsilon_F \gg kT$（$\varepsilon_F$ 为温度为 T 时的费米能，k 为玻耳兹曼常数），则电子存在的几率可近似为

$$f(\varepsilon) = \frac{1}{1+\exp[(\varepsilon - \varepsilon_F)/(kT)]} \approx \exp[-(\varepsilon - \varepsilon_F)/(kT)] \tag{5-24}$$

在能带构造中，ε_C、ε_V 分别为导带底部能级、价带顶部能级，ε_F 处于 ε_C 和 ε_V 之间（见图 5-4）。由式（5-23），导带中存在的导电电子浓度 n_e 为

$$n_e = \int_{\varepsilon_C}^{\infty} \rho_C(\varepsilon) f(\varepsilon) \mathrm{d}\varepsilon \tag{5-25}$$

式中，$\rho_C(\varepsilon)$为单位体积晶体导带中的电子状态密度，其值（参见式（3-28））为

$$\rho_C(\varepsilon)=\frac{1}{2\pi^2}\left(\frac{8\pi^2 m_e^*}{h^2}\right)^{3/2}(\varepsilon-\varepsilon_F)^{1/2} \tag{5-26}$$

式中，m_e^*为电子的有效质量；h 是普朗克常数。将式（5-24）和式（5-26）代入式（5-25）得

$$\begin{aligned} n_e &= \frac{1}{2\pi^2}\left(\frac{8\pi^2 m_e^*}{h^2}\right)\exp\left(\frac{\varepsilon_F}{kT}\right)\int_{\varepsilon_C}^{\infty}(\varepsilon-\varepsilon_F)^{1/2}\exp\left(\frac{\varepsilon}{kT}\right)\mathrm{d}\varepsilon \\ &= 2(2\pi m_e^* kT/h^2)^{3/2}\exp[-(\varepsilon_C-\varepsilon_F)/(kT)] \end{aligned} \tag{5-27}$$

由式（5-27）可以得到结论：导带里的电子数是温度和电子有效质量的函数。

本征半导体中，价带中的空穴和导带中的电子浓度相等，价带中的空穴几率 $f_h(\varepsilon)$和价带中的（不是导带中的）电子分布几率 $f(\varepsilon)$之间的关系为 $f_h(\varepsilon)=1-f(\varepsilon)$，只要 $\varepsilon_F-\varepsilon \gg kT$，便有

$$f_h(\varepsilon)=1-\frac{1}{1+\exp[(\varepsilon-\varepsilon_F)/(kT)]}\approx\exp[(\varepsilon-\varepsilon_F)/(kT)] \tag{5-28}$$

价带中的空穴浓度，可仿照导带电子浓度计算，价带中的空穴浓度 n_h 为

$$n_h=\int_{-\infty}^{\varepsilon_V}\rho_V(\varepsilon)f_h(\varepsilon)\mathrm{d}\varepsilon=2(2\pi m_h^* kT/h^2)^{3/2}\exp[-(\varepsilon_F-\varepsilon_V)/(kT)] \tag{5-29}$$

式中，$\rho_V(\varepsilon)$为单位体积晶体价带中的空穴状态密度；m_h^*为空穴的有效质量。

2. **费米能级位置**

本征半导体中，$n_e=n_h$，由式（5-27）、式（5-29）可以求出费米能级 ε_F 为

$$\varepsilon_F=\frac{1}{2}(\varepsilon_C+\varepsilon_V)+\frac{3}{4}kT\ln\frac{m_h^*}{m_e^*} \tag{5-30}$$

由于 kT 很小，而且电子和空穴的有效质量差别也不大，所以可以粗略地认为费米能级位于本征半导体的价带和导带间的正中间位置（即位于禁带中央）。

将式（5-30）代入式（5-27）、式（5-29）得到

$$\begin{aligned} n_e=n_h &= 2(2\pi kT/h^2)^{3/2}(m_e^* m_h^*)^{3/4}\exp[-(\varepsilon_C-\varepsilon_V)/(2kT)] \\ &= 2(2\pi kT/h^2)^{3/2}(m_e^* m_h^*)^{3/4}\exp[-\varepsilon_g/(2kT)] \end{aligned} \tag{5-31}$$

式中，ε_g 为导带和价带之间的能隙。

3. **本征半导体的迁移率和电阻率**

本征半导体受热后，载流子不断发生热运动，在各个方向上的数量和速度都是均布的，故不会引起宏观的迁移，也不会产生电流。但在外电场的作用下，载

流子就会有定向的漂移运动，产生电流。这种漂移运动是在杂乱无章的热运动基础上的定向运动，所以在漂移过程中，载流子不断地互相碰撞，使得大量载流子定向漂移运动的平均速度为一个恒定值，并与电场强度成正比。自由电子的自由度大，故它的迁移率较大（迁移率为单位电场强度下载流子获得的平均漂移速度）；而空穴的漂移实质上是价电子依次填补共价键上空位的结果，这种运动被约束在共价键范围内，所以空穴的自由度小，迁移率也小。在室温下，本征锗单晶中，电子的迁移率 $\mu_e = 0.39\ m^2/(V\cdot s)$，空穴的迁移率 $\mu_h = 0.19\ m^2/(V\cdot s)$；本征硅单晶中，$\mu_e = 0.14\ m^2/(V\cdot s)$，$\mu_h = 0.05\ m^2/(V\cdot s)$。硅迁移率比锗小是因其自由电子浓度较低。

在本征半导体中有电场 E，空穴将沿 E 方向作定向漂移运动，产生空穴电流密度 j_h，自由电子将逆 E 方向作定向漂移运动，产生电子电流密度 j_e，所以总电流密度 j 是两者之和，即

$$j = j_e + j_h = en_e\mu_e E + en_h\mu_h E \tag{5-32}$$

式中，e 为电子电荷的绝对值；n_e、n_h 分别为自由电子和空穴的浓度（本征半导体中 $n_e = n_h$）；E 为电场强度。所以本征半导体的电阻率 ρ 为

$$\rho = \frac{E}{j} = \frac{1}{en_e(\mu_e + \mu_h)} \tag{5-33}$$

在 300K 时，本征锗的 $\rho = 0.47\times10^{-6}\Omega\cdot m$，本征硅的 $\rho = 0.214\Omega\cdot m$，硅的 ρ 比锗的大。

本征半导体的电学特性可以归纳如下：本征激发成对地产生自由电子和空穴，所以自由电子浓度与空穴浓度相等；n_e 与能带间隙 ε_g 有近似反比关系，硅比锗的 ε_g 大，故硅比锗的 n_e 小；n_e 与温度近似正比，故温度升高时 n_e 增大；n_e 与原子密度相比是极小的，所以本征半导体的导电能力很弱。

5.2.2　杂质半导体

通常制造半导体器件的材料是杂质半导体。在本征半导体中人为地掺入五价元素或三价元素将分别获得 n 型（电子型）杂质半导体和 p 型（空穴型）杂质半导体。

1. n 型半导体

在本征半导体中掺入五价元素的杂质（磷，砷，锑）就可以使晶体中的自由电子的浓度极大地增加。这是因为五价元素的原子有五个价电子，当它顶替晶格中的一个四价元素的原子时，它的四个价电子与周围的四个硅（或锗）原子以共价键相结合后，还余下了一个价电子变成多余的，成为自由电子，如图 5-6a 所示。理论计算和实验结果表明，这个价电子能级 ε_D 非常靠近导带底，$\varepsilon_C-\varepsilon_D$ 比 ε_g 小得多（$\varepsilon_C-\varepsilon_D$ 的值在锗中掺磷为 0.012eV，在硅中掺锑为 0.039eV）。所以在常温下，

每个掺入的五价元素原子的多余价电子都具有大于 $\varepsilon_C - \varepsilon_D$ 的能量，都可以进入导带成为自由电子，因而导带中的自由电子数比本征半导体显著地增多。把这种五价元素称为施主杂质（因其能提供多余价电子），ε_D 称为施主能级，$\varepsilon_C - \varepsilon_D$ 称为施主电离能。图 5-6b、c 示出了 n 型半导体的能带图。

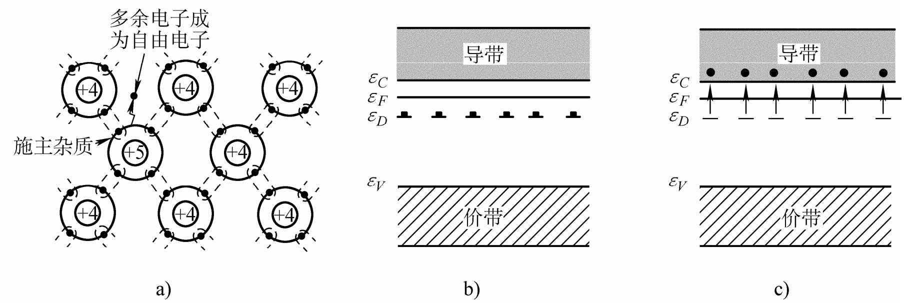


图 5-6　n 型半导体

a）纯硅掺杂晶格结构示意图　b）0K 时施主能级　c）室温时施主能级

在 n 型半导体中，由于自由电子的浓度大（1.5×10^{20} 个/m^3），故自由电子称为多数载流子，简称多子。同时由于自由电子的浓度大，由本征激发产生的空穴与它们相遇的机会也增多，故空穴复合掉的数量也增多，所以 n 型半导体中空穴的浓度（1.5×10^{12} 个/m^3）反而比本征半导体中的空穴浓度小，故把 n 型半导体中的空穴称为少数载流子，简称少子。在电场作用下，n 型半导体中的电流主要由多数载流子自由电子产生，也就是说，它是以电子导电为主，故 n 型半导体又称为电子半导体，施主杂质也称 n 型杂质。

2. p 型半导体

在本征半导体中，掺入三价元素的杂质（硼，铝，镓，铟），就可以使晶体中空穴浓度大大增加。因为三价元素的原子只有三个价电子，当它顶替晶格中的一个四价元素原子，并与周围的四个硅（或锗）原子组成四个共价键时，必然缺少一个价电子，形成一个空位置，如图 5-7a 所示。在价电子共有化运动中，相邻的四价元素原子上的价电子就很容易来填补这个空位，从而产生一个空穴。理论计算和实验结果表明，三价元素形成的允许价电子占有的能级 ε_A 非常靠近价带顶，即 $\varepsilon_A - \varepsilon_V$ 远小于 ε_g 的值（$\varepsilon_A - \varepsilon_V$ 的值在硅中掺镓为 0.065eV，掺铟为 0.16eV，锗中掺硼或铝为 0.01eV）。在常温下，处于价带中的价电子都具有大于 $\varepsilon_A - \varepsilon_V$ 的能量，都可以进入 ε_A 能级。所以每一个三价杂质元素的原子都能接受一个价电子，而在价带中产生一个空穴。我们把这种三价元素称为受主杂质（因其能接受价电子），ε_A 称为受主能级，$\varepsilon_A - \varepsilon_V$ 称为受主电离能。图 5-7b、c 示出了 p 型半导体的能带图。

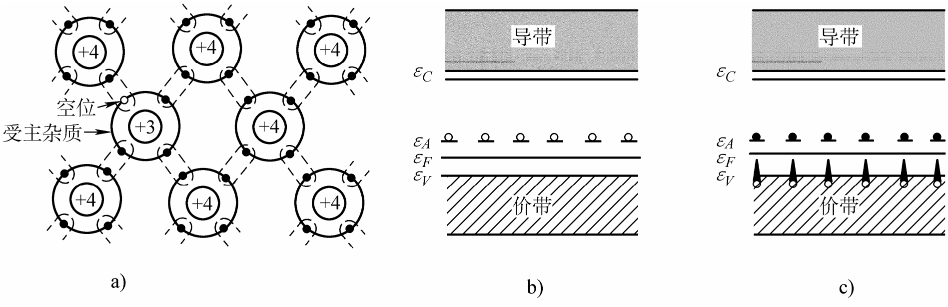


图 5-7 p 型半导体

a）纯硅掺杂晶格结构示意图 b）0K 时受主能级 c）室温时受主能级

在 p 型半导体中，因受主杂质接受价电子产生空穴的作用，使空穴浓度大大提高，故空穴为多数载流子。同时，因空穴多，本征激发的自由电子与空穴复合的机会增多，故 p 型半导体的自由电子浓度反而比本征半导体中的自由电子少，即电子是少数载流子。在电场的作用下，p 型半导体中的电流主要由多数载流子空穴产生，即它是以空穴导电为主，故 p 型半导体又称空穴型半导体，受主杂质又称 p 型杂质。

n 型半导体和 p 型半导体统称为杂质半导体，与本征半导体相比，具有如下特性：掺杂浓度与原子浓度相比虽很微小，但是却能使载流子浓度极大地提高，导电能力因而也显著地增强。掺杂浓度愈大，其导电能力也愈强；掺杂只是使一种载流子的浓度增加，因此杂质半导体主要靠多子导电。当掺入五价元素（施主杂质）时，主要靠自由电子导电。当掺入三价元素（受主杂质）时，主要靠空穴导电。

3. 杂质半导体电子和空穴数

式（5-27）、式（5-29）分别为计算本征半导体电子和空穴数的公式。在推导时并没有特殊规定，只是在能量上费米能级到导带底部的能量是 kT 的许多倍。这些同样适用于非本征半导体。问题的复杂性在于如何确定费米能级并且同时考虑施主和受主能级。条件是晶体必然是电中性，净电荷浓度必然是零。在非本征半导体中，杂质原子捐赠一个电子到导带，并留下一个正电荷，最终则有一个受主原子，接受从价带来的电子，那么便有一个负电荷，从而使其电中性公式为

$$n_e + n_A^- = n_h + n_D^+ \tag{5-34}$$

式中，n_e、n_h 分别为自由电子和空穴的浓度（单位体积的个数）；n_A^-为游离的受主原子（原子接受一个从价带来的电子）的浓度；n_D^+为游离的施主原子（原子捐赠一个电子到导带）的浓度。

怎样从实际杂质原子浓度 n_A 和 n_D 中找到游离杂质的浓度 n_A^-和 n_D^+，以图 5-6b

为例，0K 时所有施主能级皆被电子所占有，只要得到 $\varepsilon_g-\varepsilon_D$ 的能量，则可成为导电的电子。因此可由这个能量数值来计算出 n_D^+。换句话说，如果以没有占据 ε_D 能级的电子的几率乘以 n_D，便可得到 n_D^+，即

$$n_D^+ = n_D\left[1-f(\varepsilon_D)\right] \tag{5-35}$$

受主的情况类似。电子占据 ε_A 能级的几率是 $f(\varepsilon_A)$，则

$$n_A^- = n_A f(\varepsilon_A) \tag{5-36}$$

对某一确定的半导体，即已知 n_A 和 n_D、能隙 ε_g 以及电子和空穴的有效质量，把式（5-27）、式（5-29）、式（5-35）、式（5-36）表示的 n_e、n_h、n_D^+、n_A^-代入式（5-34），便可以算出 ε_F。

对 n 型半导体，$n_e \gg n_h$，$n_D^+ \gg n_A^-$，则式（5-34）简化为

$$n_e \approx n_D^+ \tag{5-37}$$

这说明，所有导电电子来自施主能级，而不是基体的晶格键合，把 n_e 及 n_D^+表达式代入式（5-37）得

$$n_C \exp[-(\varepsilon_C-\varepsilon_F)/(kT)] = n_D[1-f(\varepsilon_D)] = n_D\left\{1+\exp[(\varepsilon_F-\varepsilon_D)/(kT)]\right\}^{-1} \tag{5-38}$$

式中，$n_C = 2(2\pi m_e^* kT/h^2)^{3/2}$。计算时可以假设价带顶部的能量为零，则 $\varepsilon_F-\varepsilon_C=\varepsilon_F-\varepsilon_g$，式（5-38）即为

$$n_C \exp[(\varepsilon_F-\varepsilon_g)/(kT)] = n_D\left\{1+\exp[(\varepsilon_F-\varepsilon_D)/(kT)]\right\}^{-1} \tag{5-39}$$

由式（5-39）得到

$$\varepsilon_F = \varepsilon_D + kT\ln\left\{-\frac{1}{2}+\frac{1}{2}\left[1+4\frac{n_D}{n_C}\exp\left(\frac{\varepsilon_g-\varepsilon_D}{kT}\right)\right]^{1/2}\right\} \tag{5-40}$$

从式（5-40）可见，ε_F 是温度的函数。现分三个温度区域讨论式（5-40）的物理意义。

① 在低温时，$\exp[(\varepsilon_g-\varepsilon_D)/(kT)] \gg 1$，则

$$\varepsilon_F = (\varepsilon_g+\varepsilon_D)/2 + (kT/2)\ln(n_D/n_C) \tag{5-41}$$

导电电子浓度为

$$n_e = \sqrt{n_C n_D}\exp[-(\varepsilon_g-\varepsilon_D)/(kT)] \tag{5-42}$$

式（5-41）说明费米能级位于导带底和施主能级之间（$\ln(n_D/n_C)<0$）。

② 当温度升高至 $\exp[(\varepsilon_g-\varepsilon_D)/(kT)]\approx 1$ 时，则

$$\varepsilon_F = \varepsilon_D - kT\ln(n_C/n_D) \tag{5-43}$$

导电电子浓度为

$$n_e = n_D \tag{5-44}$$

式（5-43）说明费米能级随温度升高，不断向本征半导体费米能级接近。因为在这个区域施主能级的电子全部跃迁至导带。

③ 更高的温度区域时，半导体成为本征半导体。此区域即为本征区域。费米能级近似位于禁带中央。

图 5-8a 表示上述分析结果的物理图像。

对 p 型半导体也可以做类似的讨论，结果表示在图 5-8b 中。图 5-8c 说明杂质含量对非本征半导体费米能级位置的影响。如图 5-8c 所示，对于 n 型半导体，当施主浓度为 $10^{26}\ m^{-3}$ 时，费米能级在 600℃以下靠近导带的底部。随着施主浓度的减少，费米能级向比导带更低的能量方向移动；对于 p 型半导体，当受主浓度为 $10^{25}\ m^{-3}$ 时，费米能级在 400℃以下位于接近价带的顶部，随着受主浓度减少，费米能级向价带顶部更高能量方向移动。

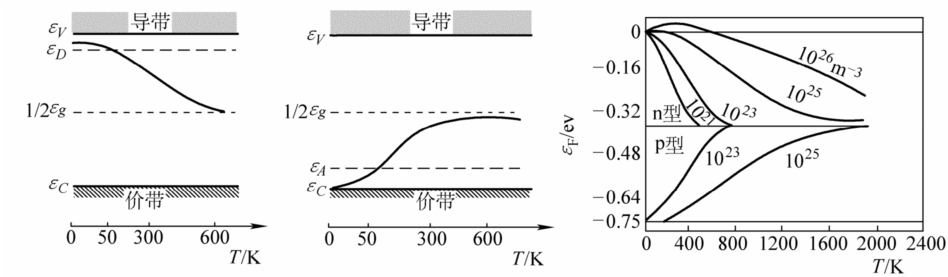


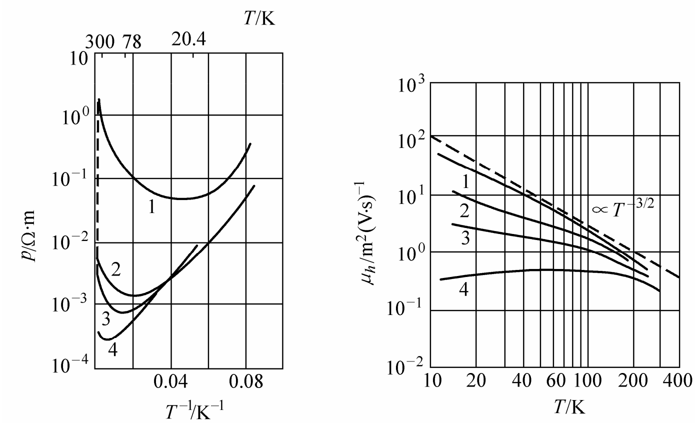
图 5-8 费米能随温度的变化

a）n 型半导体 b）p 型半导体 c）不同掺杂浓度的 n 型和 p 型 Ge 半导体

4. 杂质半导体迁移率与温度的关系

图 5-9a 所示为测得的含不同杂质浓度 Ge 半导体电阻率 ρ 与温度 T 的关系曲线。已知其各温度下的霍耳常数，那么便可绘出 Ge 的空穴迁移率 μ_h 与温度 T 的关系，如图 5-9b 所示。

载流子由于晶格的热振动而被散射，迁移率 μ 与温度 T 具有 $T^{-3/2}$ 的关系。离子化的杂质同样散射载流子，杂质含量愈高，迁移率愈低。图 5-9b 说明在高温或者低杂质浓度时晶格散射起主要作用；当杂质浓度高时，杂质散射起主要作用。



a）　　　　b）

图 5-9　Ge 电性能与温度的关系

a）电阻率 ρ 与温度 T 的关系　b）空穴迁移率 μ_h 与温度 T 的关系

杂质浓度（n_D–n_A）/ m^3：1—1.0×10^{19}　2—1.7×10^{21}　3—7.5×10^{21}　4—5.5×10^{22}

5.2.3　半导体的表面能级

前面所述的能带结构是在无限扩展的三维晶体产生的周期场的前提下得到的。在材料的表面，势场不再与晶体内部的周期性势场相同，所以材料表面的电子能级分布就会发生变化。图 5-10 表示晶体表面的能带结构。此时，位于晶体表面的电子能够占据禁带内的能级。如果这些能级上有电子占据，也会对晶体内部的能带产生影响，从而使能带发生弯曲。

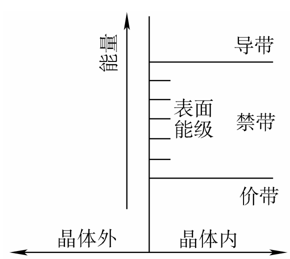


图 5-10　晶体表面的能带结构

图 5-11 表示 n 型半导体的表面能级模型。判断一个系统是否处于平衡状态的根据是看其费米能级是否相等。两个分立的材料，费米面可以不一样。但如果这两个材料连成一个系统，就会在这两个材料之间发生电荷的移动，最终使费米能级相等。如果像图 5-11a 中那样，位于表面的电子只填充到 0 的位置，那么表面的费米能级和内部的费米能级不相等，所以不是平衡状态。为了达到平衡状态，位于表面附近的电子就会移到表面去，占据表面电子能级，最后表面的费米能级与内部的费米能级相等。因为能带是连续性的，而且禁带宽度有不可改变性，所以形成了如图 5-11b 所示的表面能带结构。由于电子从内部向表面迁移，在表面会出现负电荷，而接近表面的内部会因缺少电子而出现带正电荷的空穴。这些空穴的存在，使 n 型半导体的表面附近出现了一个 p 型的反转层。同样，在 p 型半

导体的表面附近出现的是 n 型反转层。在 p 型半导体中，激发电子向半导体表面移动，而空穴向半导体内部移动。

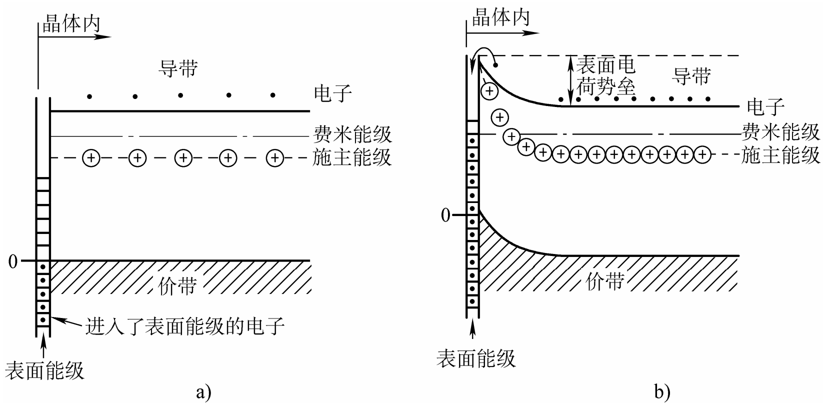


图 5-11　n 型半导体的表面能级

a）平衡前　b）平衡后

5.2.4　半导体接触

在一个半导体中形成 p 型和 n 型两个区域，p 型和 n 型的界面就是 pn 结。pn 结具有单向导电性，是许多重要半导体器件的核心。

掺杂浓度较低的半导体与金属所形成的接触和 pn 结一样具有单向导电性（此时金属功函数 > 半导体功函数，材料的功函数是指材料的费米能级与真空能级之间的差值），其频率特性优于 pn 结。高掺杂半导体与金属形成的接触通常为低阻非整流接触（此时金属功函数 < 半导体功函数），几乎所有的电子器件中都利用它来实现欧姆接触。

1. 金属—半导体结

典型的金属与半导体接触有两类：一类是整流接触，即制成肖特基势垒二极管，另一类是非整流接触，即欧姆接触。当一块金属和一块半导体接触时，能量达到平衡时，其费米能级一定要连续。图 5-12 示意了 n 型半导体与不同功函数的金属接触成结时能带结构变化的情况。假设金属功函数为 Φ_m，半导体的功函数为 Φ_s。当 $\Phi_m > \Phi_s$ 时，电子从半导体流入金属，直到平衡。此时费米能级连续通过结，这是由于接近结的半导体中的电子，离开施主进入金属的结果。这个过程在半导体接近结的部分形成耗尽层。此时离子化的施主形成了正的空间电荷区，金属一边接触面形成负的表面电荷。耗尽区（或称势垒区，或者空间电荷区）之所以这样称呼它是因为它已经耗尽自由载流子，结果电阻很高。电子继续流动，直到双电层建立了足够强的电场，阻止电子继续流动，从而造成了在半导体一边能带向

上弯，形成一个势垒为 $\Phi_m-\Phi_s$（图 5-12a）。此时由于热扰动，只有极少数电子从金属流向半导体而且半导体导带里只有极少的电子有足够能量越过势垒进入金属，在平衡时这二类电流相等，没有净电流通过结。另一方面，如果 $\Phi_m<\Phi_s$，情况完全不同。电子将从金属流向半导体，引起表面电荷积累，直到平衡时费米能级连续通过结。最终内部自发形成的电场导致半导体的能带下降弯曲。在这种情况下，结处无势垒，电子通过结“来去”自由，这种结称为欧姆结（图 5-12b）。

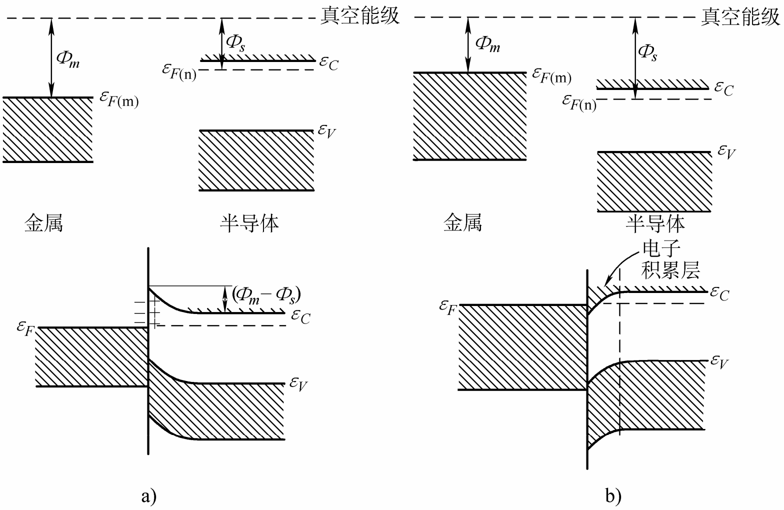


图 5-12　金属—n 型半导体成结前后的能带结构

a）$\Phi_m>\Phi_s$ 整流结　b）$\Phi_m<\Phi_s$ 欧姆结

类似的论证可以说明，当 p 型半导体与金属接触成结时，当 $\Phi_m<\Phi_s$ 时形成整流结，而 $\Phi_m>\Phi_s$ 时形成欧姆结。

对 n 型半导体的整流结，如果加上偏压 V，半导体边与正极连接，金属边与负极连接，则半导体能带以 eV 值下降，这就导致了势垒进一步升高（势垒高度为 $|eV|+|\Phi_m-\Phi_s|$），并使耗尽层宽度增加（图 5-13b）。这种外加电压的条件称为加反向偏压。在这种条件下，从半导体流向金属的电子减少，因而势垒高度增大。但是电子从金属流向半导体并没有改变，那么此时的反向电流在一个低的 V 值就达到饱和，并且基本上与 V 无关。当外加电压极性改变，此时加的就是正偏压（图 5-13c），那么阻碍电子从半导体向金属流动的势垒下降了 $|eV|$。可以证明，此时的电流与正偏压成指数关系增加。这就是单向导电性，反向电流很小。单向导电称为整流，通常称肖特基整流。

上述讨论的情况是不考虑半导体表面能级的理想状态。不过，由于常用的半导体材料的表面能级密度很高，与金属接触都会形成势垒。所以实际工作中，常

通过重掺杂半导体与金属接触，使其势垒很薄，电子可以通过隧道效应穿过势垒，从而形成欧姆接触。

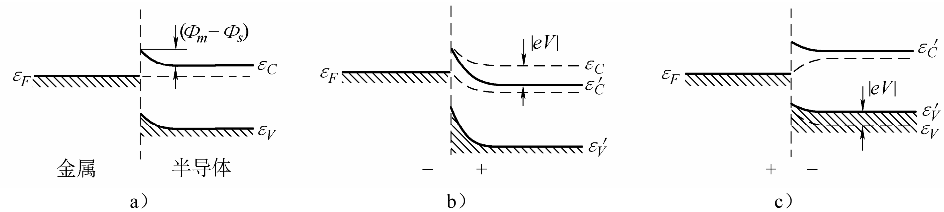


图 5-13　金属—n 型半导体结的偏压效应（$\Phi_m > \Phi_s$）

a）零偏压　b）反向偏压　c）正向偏压

肖特基势垒二极管高频特性好，开关速度快，由于它是杂质引起的多数载流子在起作用，不是由于热产生的本征激发的少数载流子起作用，所以热噪声很低。

欧姆接触是设计和制造超高频、大功率器件的关键问题，因为半导体元件都需要通过电极引线与外部电路进行电学连接，而欧姆接触效应则广泛地应用于这些电极引线的设计生产中。

2. pn 结

pn 结是由单晶半导体上相邻的两个区——p 型区和 n 型区构成的。n 区掺施主杂质，浓度为 n_D；p 区掺受主杂质，浓度为 n_A。为了便于理解，可以假想 pn 结是由同一种材料进行不同掺杂的两块样品紧密接触而形成的。那么在接触前，在人们所关心的温度和掺杂范围内，n 型区的费米能级高于 p 型区的费米能级（参见图 5-8，可认为接触前 p 型和 n 型的导带底部能量相等），在 n 型区存在自由电子以及相等浓度的正电离施主杂质原子，以保持电中性。在 p 型区存在自由空穴以及相同浓度的负电离受主杂质原子，以保持电中性。于是 n 型区的多数载流子是电子，p 型区的多数载流子是空穴。在两区紧密接触后，一开始，由于两区的费米能级不同，电子和空穴的浓度不同，n 区的电子通过扩散进入 p 区的未被电子占据态，降低了电子的能量。p 区空穴也要向 n 区扩散。这种电荷转移的结果是在 n 区出现电离施主构成的正空间电荷区，而在 p 区出现电离受主构成的空间负电荷区，因此空间电荷区内就形成了从 n 区指向 p 区的电场，称为内建电场（图 5-14）。它倾向于阻止 n 区电子和 p 区空穴越过分界面向对方扩散，换句话说，两区的能级发生了相对位移，最后当能级的相对移动使两侧费米能级拉平时，动平衡就达到了。动态平衡的条件是

$$e\Phi = \varepsilon_{F(\mathrm{n})} - \varepsilon_{F(\mathrm{p})} \tag{5-45}$$

式中，Φ 是接触电势差；$\varepsilon_{F(\mathrm{n})}$和 $\varepsilon_{F(\mathrm{p})}$分别为未接触前 n 型区和 p 型区的费米能级；$e$ 为电子电荷绝对值。

pn 结最令人感兴趣的特征是，在平衡条件下，结的两侧存在式（5-45）给出

的接触电势差，结果使 n 型区的电势比 p 型区高 Φ。对电子而言，这意味着 p 型区一侧势能高于 n 型区势能，结区形成了电子势垒，其作用是阻碍电子向 p 区扩散。同理，对从 p 型区向 n 型区扩散的空穴而言，结区也是势垒。势垒的高度为 $e\Phi$。在空间电荷区内自由载流子数目很少，称为耗尽层，如图 5-14b 中所示的$-x_p$至 x_n 区域，显然该层是高阻层。耗尽层的宽度可以从泊松方程中求得。

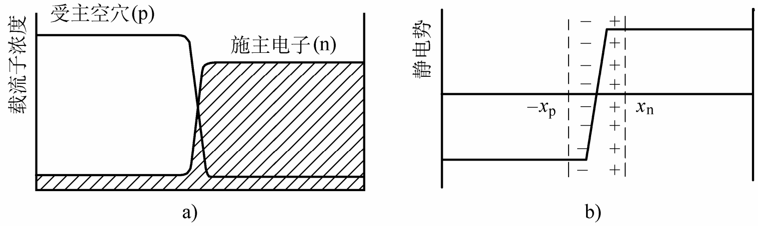


图 5-14　pn 结

a）无偏压时结区的空穴和电子浓度变化　b）受主负离子和施主正离子产生的内电场，从 x_p 至 x_n 称为耗尽层

pn 结可以起整流作用。如果在结上施加某个方向的电压（p 端接电源正极，n 端接电源负极），将会有较大电流流动。但如果电压反向，那就只有很小的电流流动。如果在结上施加一个交流电压，那么电流主要向一个方向流动，这就是 pn 结的整流作用。pn 结整流的原理如下所述。

原来 p 区相对于 n 区的电势为$-\Phi$，如果加电压 V_0 于 p 区，则 p 区相对于 n 区的电势改为 $-(\Phi-V_0)$，这时势垒高度为 $e(\Phi-V_0)$。如果 V_0 为正电压，则能带图中势垒将降低，如图 5-15 所示。在这种情况下，势垒不再能完全抵消电子和空穴的扩散作用。用 p 和 n 表示半导体类型，下标 0 特指热平衡情况，因此，n_{n0} 和 n_{p0} 分别是热平衡下 n 区和 p 区的电子浓度，p_{n0} 和 p_{p0} 是热平衡下 n 区和 p 区的空穴浓度。

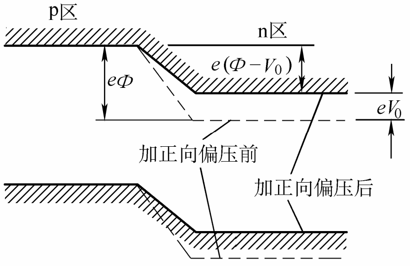


图 5-15　正向偏压引起的 pn 结能带的改变

根据式（5-27）可知，在平衡条件下，电子浓度在 n 区和 p 区之比为

$$\frac{n_{n0}}{n_{p0}}=\exp\left(\frac{e\Phi}{kT}\right) \tag{5-46}$$

同理，根据式（5-29）可知，在平衡条件下，空穴浓度在 p 区和 n 区之比为

$$\frac{p_{\mathrm{p0}}}{p_{\mathrm{n0}}}=\exp\left(\frac{e\Phi}{kT}\right) \tag{5-47}$$

从上式可知，在热平衡条件下，在结两边耗尽层边界电子与空穴的浓度是与电势差 Φ 相关的，显然当施加外电压 V_0 而引起静电势改变时，也应有同样的表达式。此时式（5-46）修正为

$$n_{\mathrm{n}}=n_{\mathrm{p}}\exp\left[\frac{e(\Phi-V_0)}{kT}\right] \tag{5-48}$$

式中，n_{n} 和 n_{p} 分别是在 n 区和 p 区的耗尽层边界上非平衡电子浓度；V_0 为正时为正向偏置电压，V_0 为负时为反向偏压。在低注入条件下，注入的少数载流子浓度比多数载流子浓度小得多，因此 $n_{\mathrm{n}}\approx n_{\mathrm{n0}}$，将该条件与式（5-46）代入式（5-48），得到了 p 区耗尽层边界（$x=-x_{\mathrm{p}}$，参见图 5-14b）的电子浓度为

$$n_{\mathrm{p}}=n_{\mathrm{p0}}\exp\left(\frac{eV_0}{kT}\right) \tag{5-49}$$

即

$$n_{\mathrm{p}}-n_{\mathrm{p0}}=n_{\mathrm{p0}}\left[\exp\left(\frac{eV_0}{kT}\right)-1\right] \tag{5-50}$$

同理有 n 区耗尽层边界（$x=x_{\mathrm{n}}$，参见图 5-14b）的空穴浓度为

$$p_{\mathrm{n}}-p_{\mathrm{n0}}=p_{\mathrm{n0}}\left[\exp\left(\frac{eV_0}{kT}\right)-1\right] \tag{5-51}$$

从式(5-51)可见，由于外电压存在，注入的空穴不断向 n 区扩散，扩散流为 $D_p\dfrac{\mathrm{d}p_{\mathrm{n}}}{\mathrm{d}x}$，其中 D_p 表示空穴扩散系数。但由于它们是多余少数载流子，在扩散过程中又不断复合而消失。根据连续性方程，在稳态情况下，由扩散流不均匀而造成的积累率与非平衡载流子的净复合率相等，向 n 区扩散的空穴的积累率为 $\dfrac{\mathrm{d}}{\mathrm{d}x}\left(D_p\dfrac{\mathrm{d}p_{\mathrm{n}}}{\mathrm{d}x}\right)$。净复合率 U 与多余少数载流子浓度成正比，比例常数为 $1/\tau_p$，τ_p 是描述多余少数载流子平均存在的时间，称为多余少数载流子寿命。在热平衡下 U=0。净复合率 U 与多余少数载流子浓度的关系为

$$U=\frac{p_{\mathrm{n}}-p_{\mathrm{n0}}}{\tau_p} \tag{5-52}$$

利用式（5-52）可得

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(D_p\frac{\mathrm{d}p_{\mathrm{n}}}{\mathrm{d}x}\right)=\frac{p_{\mathrm{n}}-p_{\mathrm{n0}}}{\tau_p} \tag{5-53}$$

即

$$\frac{d^2 p_n}{dx^2} - \frac{p_n - p_{n0}}{D_p \tau_p} = 0 \tag{5-54}$$

方程式（5-54）的边界条件为

$$\left.\begin{array}{l} x = x_n 时, p_n - p_{n0} = p_{n0}\left[\exp\left(\dfrac{eV_0}{kT}\right) - 1\right] \\ x \to \infty 时, p_n = p_{n0} \end{array}\right\} \tag{5-55}$$

则方程式（5-54）的解为

$$p_n - p_{n0} = p_{n0}\left[\exp\left(\frac{eV_0}{kT}\right) - 1\right]\exp\left(-\frac{x - x_n}{L_p}\right) \tag{5-56}$$

式中，$L_p = \sqrt{D_p \tau_p}$，是空穴在 n 区的扩散长度。在 $x = x_n$ 时，由空穴产生的电流密度 $J_p(x_n)$为

$$J_p(x_n) = -eD_p \left.\frac{dp_n}{dx}\right|_{x_n} = \frac{eD_p}{L_p} p_{n0}\left[\exp\left(\frac{eV_0}{kT}\right) - 1\right] \tag{5-57}$$

同理，可获得进入 p 区的电流密度 $J_n(-x_p)$为

$$J_n(-x_p) = -eD_n \left.\frac{dn_p}{dx}\right|_{-x_p} = \frac{eD_n}{L_n} n_{p0}\left[\exp\left(\frac{eV_0}{kT}\right) - 1\right] \tag{5-58}$$

这里，D_n 表示电子扩散系数；L_n 是电子的扩散长度。

因此，通过 pn 结界面的总电流 J 为

$$J = J_p(x_n) + J_n(-x_p) = J_0\left[\exp\left(\frac{eV_0}{kT}\right) - 1\right] \tag{5-59}$$

其中，$J_0 = \dfrac{eD_p p_{n0}}{L_p} + \dfrac{eD_n n_{p0}}{L_n}$。

由式（5-59）可以看出，在加正向偏压时，电流基本上按 $\exp\left(\dfrac{eV_0}{kT}\right)$ 指数增加，在加反向偏压时，电流密度很快降到饱和值$-J_0$。由此可以达到整流效应。

3. 结晶体管和场效应管

结晶体管（俗称三极管），其主要功能是放大。图 5-16a、b 分别为 pnp 晶体管的共基极接线图和电子能带结构图。这种晶体管，中部是基极区，很薄，而且相对于发射极区的掺杂浓度也是很低的。其工作状态取决于多子和少子的相互作用。由于该器件涉及两种载流子，因此常把它归类为双极晶体管。图 5-17 所示为 npn 型晶体管共基极接线时的能带结构及载流子流动情况。

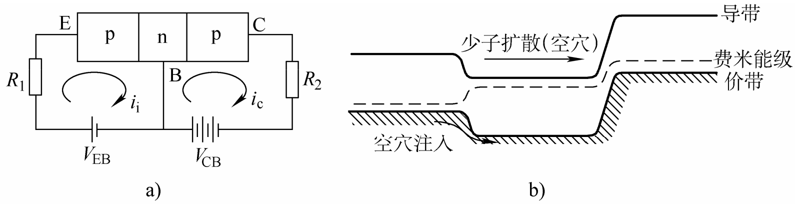


图 5-16 pnp 型结晶体管共基极接线（a）及电子能带结构图（b）

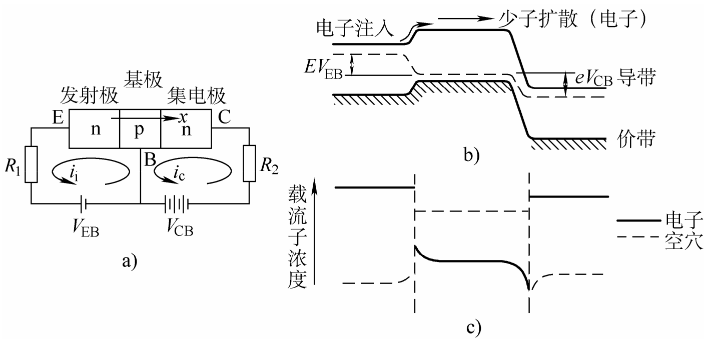
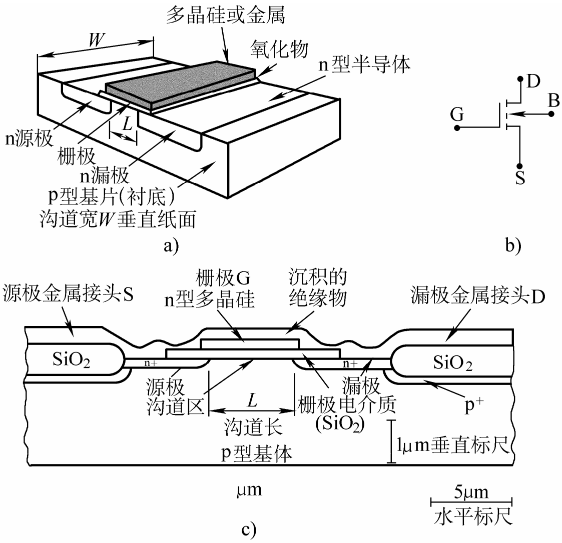


图 5-17 npn 型共基极晶体管的工作原理示意图

a）接线图 b）能带结构图 c）载流子浓度分布图

场效应管是另一种类型的晶体管，它也是集成电路中的基本器件。三极管和场效应管的基本差别在于：场效应管电流是仅靠多子输运形成的，多子流是控制电流，因此降低了对环境温度变化的敏感，以及其他外部因素的影响。场效应管有几种形式，主要有结场效应管和金属氧化物半导体场效应管。使用最多的是金属氧化物 n 型半导体场效应管，英文名字为 n－MOSFET 或 NMOS，图 5-18 示意了它的结构。由图 5-18 可见，在 p 型硅基片上，制成两个孤立的 n 型硅区域。在这种器件中，电子进入的触点称源极，电子离开点为漏极，在 n 型硅的源和漏之间为 p 型硅区域，就在该 p 型硅的表面沉积一层 SiO_2，作为绝缘层。在 SiO_2 的上面再沉积一层多晶硅或金属，以形成该晶体管的第三个接触点，它被称为栅极。因为 SiO_2 是良好的绝缘体，因此栅极与它下面的 p 型硅电路上并不直接接触。对于简化的 NMOS 器件，当没有电压加到栅极上时，栅极下面的 p 型硅多数载流子是空穴，只有极少的电子被吸入漏极。当栅极加上正向电压时，这个电场从临近的 n^+源极和正好在栅极下面的也就是 SiO_2 薄层下面的漏区域吸引电子，以至于使这个区域成为 n 型硅（图 5-19）。当电子在该沟道内存在时，在源极和漏极之间存在导电通路。即存在正向电位差时，源极和漏极之间有电流流过。MOS 场效应管同样具有放大功能，不过测量的是电压而不是电流。



a)　b)　c)

图 5-18　NMOS 结构示意图

a）结构全貌　b）N 沟道增强型 MOS 结构符号　c）结构横截面剖面图

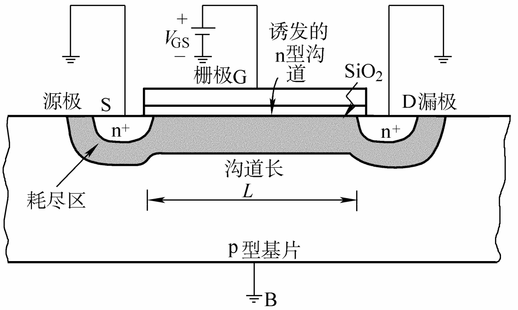


图 5-19　具有正门电压 V_{GS} 的理想 NMOS 器件

5.3　超导体

在前面的章节中，我们基本上是利用独立电子近似处理电子体系，把多电子问题简化为单电子问题。实际上电子彼此之间存在各种直接或间接的相互作用与关联，虽然这些相互作用细节在研究能带结构时并不重要，但是在另一些现象中却可能占据主导地位。

在许多金属中，在极低温度下，会出现某种电子有序态，这称为超导态，在这种状态中传导电子所表现的一些宏观性质不同于理想费米气，实质上它是具有相互作用的费米系统在低温下的一个宏观量子现象。超导电性并不是个别金属所

特有的，至少二十余种金属元素在极低温度下能够成为超导体，甚至某些半导体、多元金属氧化物、一系列合金在适当条件下也可处于超导态。

5.3.1 超导态三个重要特征

1. 不存在可观测的直流电阻

昂尼斯（K. Onnes）等人曾进行过这样的实验，先将超导体做成的环放入磁场中，此时温度 $T>T_c$，环中无电流，然后再将环冷却至 T_c 以下，使环变成超导态（T_c 即为超导体的临界转变温度），此时环中仍无电流；但若突然去掉磁场，则环内将有感应电流产生。这是由于电磁感应作用的结果，按楞次定律，该电流应沿反抗磁通变化的方向流动。如果此环的电阻确实为零，那么这个电流就应长期无损地流下去。事实上经过长达几年的观察，没发现电流有任何衰减，这就有力地证明了超导体的电阻确实为零，是完全导电性的。同时也说明了超导体是等电位的，即超导体内没有电场。

2. 完全的抗磁性

迈斯纳（W. Meissner）等人由下列实验分析了超导体在外磁场中的特性。先将超导体冷却至超导态（$T<T_c$），然后加磁场，发现磁场不能进入超导体内。若在常温下将超导体先放入磁场内，则有磁力线穿过超导体；然后再将超导体冷却至 T_c 以下，发现磁场从超导体内被排出，即超导体内无磁场，磁感应强度 B 为零。这一现象说明了超导体具有完全的抗磁性，又称迈斯纳效应。

超导态为什么会出现完全抗磁性呢？这是由于外磁场在试样表面感应产生一个感应电流，如图 5-20b 示。此电流由于所经路径电阻为零，故它所产生的附加磁场总是与外磁场大小相等，方向相反，因而使超导体内的合成磁场为零。由于此感应电流能将外磁场从超导体内挤出（图 5-20c），故称抗磁感应电流，又因其能起着屏蔽磁场的作用，又称屏蔽电流。

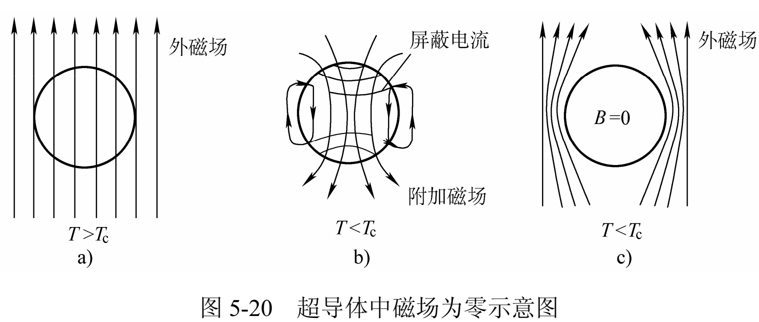


图 5-20 超导体中磁场为零示意图

a）$T>T_c$ 时外磁场穿过超导体 b）感应电流产生附加磁场 c）外磁场从超导体内挤出

3. 超导体存在超导能隙

一般来说超导体存在超导能隙。金属处于正常态时，基态与最低激发态之间

没有能隙，一旦发生了超导转变，能隙就出现。因此，超导体中的能隙是与相互作用的电子气相联系的，图 5-21 给出了超导态中的能隙。处于能隙以上激发态的电子其行为如同正常电子在射频场中一样：这些电子产生电阻，在直流情况下这些电子为超导电子所短路。在绝对零度，能隙以上没有电子。设想该能隙是以费米能级为中心，间隔为 $\varepsilon_g = 2\varDelta$。应该注意，超导体的能隙同绝缘体的能隙相比较具有全然不同的本质。

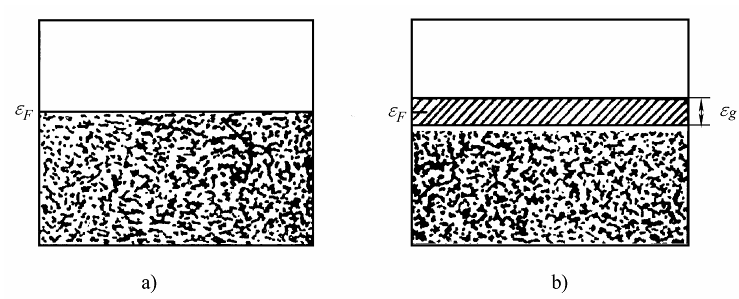


图 5-21　正常态中的导带（图 a）和超导态中费米能级附近的能隙（图 b）

超导体比热容、电子隧道效应、与频率相关的电磁行为和声衰减这些现象表现出能隙存在。

正常金属的低温比热容具有 $AT+BT^3$ 形式（A、B 为常数，T 为温度），其线性项是由电子激发引起的，立方项是由晶格振动引起的。实验发现，在 $T \leqslant T_c$ 的零磁场下，超导体比热容根本改变了上述形式。首先在 $T=T_c$ 时，其比热容跳到一个较高的值，然后随温度下降到低于正常金属的相应值。利用外磁场将金属转变成正常态，对临界温度以下的超导态与正常态的比热容进行比较，如图 5-22 所示。从该图的分析可知，在超导态中，电子对比热容的线性贡献已经被指数形式 $\exp[-\varDelta/(kT)]$ 代替。这是激发态与基态之间存在能隙的特征热行为。能隙存在是超导态的一个特征，但并非普适性的表现。对于低温超导体，理论和实验都表明能隙 $2\varDelta$ 具有 kT_c 量级。

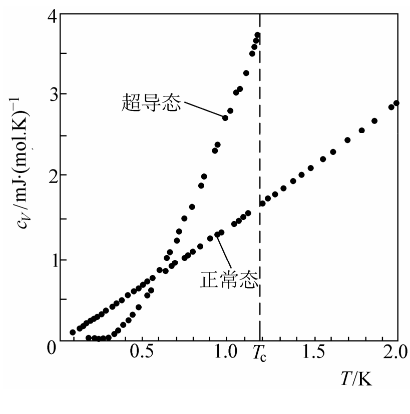


图 5-22　金属铝处于正常态和超导态时，其低温比热容的比较（正常态是通过外加强磁场得到的）

考虑被绝缘体薄层隔开的两块金属，如果绝缘层足够薄，由于量子隧道效应的存在，电子将具有一定的几率贯穿阻挡层。在热平衡时，足够的电子将从一块金属渡越到另一块金属，使两块金属的化学势相等。当两块金属都处于正常态时，外加电压将引起一块金属的化学势高于另一块，更多的电子穿过

绝缘层势垒，因此，正常态金属结的“隧道电流”应遵循欧姆定律，电流正比于外电压。当两块金属之一处于超导态时，在远低于临界温度下，直到外电压 V 达到阈值电压，$eV = \Delta$时，才有电流产生。Δ值刚好与低温比热容测量的相应值一致，这确信了超导体能隙的存在。当温度升高趋于 T_c 时，阈值电压减少，这表明能隙本身随温度增加而减少。

金属对电磁幅射的响应是由电导对频率依赖关系确定的。由于在超导体中电子系统的激发谱是由能隙Δ表征的，可以预料在频率比$\Delta/\hbar$ 小时，其交流电导和正常态有显著差别，但在频率比$\Delta/\hbar$ 大时，两者应基本相同。除了相当接近 T_c 的温度区外，典型的$\Delta/\hbar$ 值是微波和红外频谱区。由此可知，在光学频率区观察到的交流行为和正常态行为无区别。在红外区将出现与正常态的偏离，而且微波频率区将完全表现出由于能隙存在电子不能被激发的特征。

当声波在金属中传播时，由于离子位移所产生的微波电场能够把能量给予费米面附近的电子，从声波中取走能量。在远低于 T_c 温度范围，当 $\omega\hbar < 2\Delta$ 时，超导态中的声衰减显著低于正常态金属。

以上几种表现形式，在大部分超导体中已经被观察到。

5.3.2 超导态三个性能指标

评价实用超导材料有三个性能指标。其一是超导体的临界转变温度 T_c，它愈接近室温愈好，以便于利用。目前 T_c 最高的金属氧化物高温超导体其 T_c 也只有 140K 左右。

第二个指标是临界磁场强度 B_c。温度 $T < T_c$ 时，将磁场作用于超导体，当磁场强度大于 B_c 时，磁力线将穿入超导体，即磁场破坏了超导态，使超导体回到了正常态，此时的磁场强度称为临界磁场强度。很明显，B_c 大小和温度 T 是有关系的，它们之间的关系为

$$B_c = B_{c0}\left[1-\left(\frac{T}{T_c}\right)^2\right] \tag{5-60}$$

式中，B_{c0} 是 0K 时超导体的临界磁场。因此，可以定义临界磁场就是破坏超导态的最小磁场。B_c 与材料性质有关，不同超导材料临界磁场变化范围很大。

临界电流密度是评价超导体的第三个指标，除磁场影响超导转变温度外，通过的电流密度也会对超导态起影响作用。若把温度 T 从超导转变温度下降，则超导体的临界磁场也随之增加。如果输入电流所产生的磁场与外加磁场之和超过超导体的临界磁场 B_c 时，则超导态被破坏。此时通过的电流密度称为临界电流密度 J_c。随着外磁场的增加，J_c 必须相应减小，从而保持超导态。故临界电流密度是保持超导态的最大输入电流。

5.3.3　两类超导体

超导体分为两类。元素超导体除 V、Nb、Ta 以外都属于Ⅰ类超导体，它们的磁化行为如图 5-23 所示。图中 M 代表磁化强度（负号表示为抗磁性），B_a 代表外磁场，B_c 为临界磁场强度，T_c 为超导转变温度。V、Nb、Ta 以及合金和化合物超导体都是Ⅱ类超导体，它们的磁化行为如图 5-24 所示。由图 5-24 可见，它们有两个临界磁场：即下临界磁场 B_{c1} 和上临界磁场 B_{c2}。在温度低于 T_c 条件下，外磁场 $B_a < B_{c1}$ 时，Ⅱ类超导体像Ⅰ类超导体那样，都处于迈斯纳状态（$B = 0$）。当外磁场 B_a 介于 B_{c1} 和 B_{c2} 之间时，Ⅱ类超导体处于混合态（也称涡旋态）。其中部分区域有磁感应线穿过，属于正常态，它的周围却是超导态。但超导体仍具有零电阻特性。当外磁场 B_a 达到 B_{c2} 时，正常态数目增多到彼此相接触，整体超导体都变成了正常态。Ⅱ类超导体按其磁化曲线是否可逆，又分为理想的和非理想的Ⅱ类超导体。

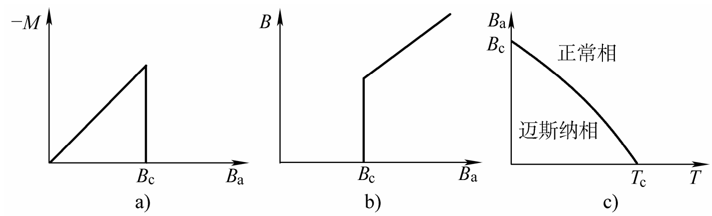


图 5-23　Ⅰ类超导体的磁化行为和几个状态（相）界线

a）$-M$ 与 B_a 关系　b）B 与 B_a 关系　c）B_a 与 T 关系

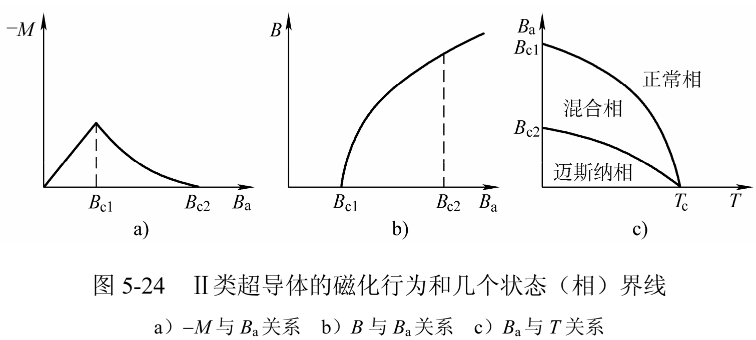


图 5-24　Ⅱ类超导体的磁化行为和几个状态（相）界线

a）$-M$ 与 B_a 关系　b）B 与 B_a 关系　c）B_a 与 T 关系

5.3.4　超导现象的物理本质

超导的微观物理本质终于在超导现象发现后的 46 年，即 1957 年由巴丁（J. Bardeen）、库柏（L. N. Cooper）和施瑞弗（J. R. Sehriffer）等人发现，简称为 BCS 理论。这个理论认为，超导现象产生的原因是由于超导体中的电子在超导态时，

电子之间存在着特殊的吸引力，而不是正常态时电子之间的静电斥力。这种吸引力使电子双双结成电子对，它是超导态电子与晶格点阵间相互作用产生的结果。如图 2-25 所示，当超导体内处于超导态的某一电子 e_1 在晶体中运动时，它周围的正离子点阵将被这个电子吸引，并向该电子靠拢以降低静电能，从而使此局部区域的正电荷密度增加，而这个带正电的区域又会对临近电子 e_2 产生吸引力，正是由于这种吸引力克服了静电斥力，使动量和自旋方向相反的两个电子 e_1、e_2 结成了电子对，称为库柏电子对。

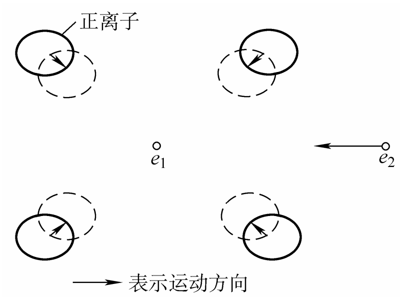


图 5-25 电子与正离子相互作用形成电子对示意图

显而易见，组成库柏对的电子 e_1 和 e_2 之间的这种相互吸引作用与正离子的振动有关，而且在超导体内，这些正离子的运动是相互牵连的，某个正离子的振动，会使邻近正离子也发生振动，一个个传下去，其结果是形成了一个以声速在晶格上传播的波动，即格波。

据理论计算，对能量相近似的两个电子，由晶格引起的这种间接作用力是吸引力。显然电子与晶格间作用越强，这种吸引力就越大。而且根据量子统计法则，如果每对电子的总动量都相等，那么每一对中的两个电子之间的吸引力也大大加强。因此，在电子结成库柏电子对时，每对电子的总动量都是相当的。库柏对中两个电子的相互作用范围为 $10^{-9}\sim10^{-6}$m，而一般晶格中原子之间的距离只有 10^{-10}m，由此看出，互相吸引而结成对的两个电子相距可能很远，这是因为电子是通过格波而在相互作用的。

材料变为超导态后，由于电子结成库柏对，使能量降低而成为一种稳定态。一个超导电子对的能量比形成它的单独的两个正常态的电子的能量低 2Δ，这个降低的能量 2Δ 称为超导体的能隙，而正常态电子则处于能隙以上的能量更高的状态。能隙的大小与温度有关，有关系式

$$2\Delta = 6.4kT_c\left[1-\left(\frac{T}{T_c}\right)\right]^{1/2} \tag{5-61}$$

式中，k 为玻耳兹曼常数，T_c 为由正常态转变为超导态的临界温度。由式可见，$T=0$ 时，能隙最大。当电子对获得的能量大于 2Δ 时就进入正常态，即电子对被拆开成两个独立的正常态电子。当温度或外磁场强度增加时，电子对获得能量，能隙就减小。当温度增加到 $T=T_c$，外磁场强度增加到 $H=H_c$ 时，能隙减小到零，电子对全部被拆开成正常态电子，于是材料即由超导态转变为正常态。由此可知，为什么温度越低，超导体就越稳定。

前已提及，超导态的电子对有一基本特性，即每个电子对在运动中的总动量保持不变，故在通以直流电时，超导体中的电子对将无阻力地通过晶格运动。这是因为任何时候，晶格（缺陷）散射电子对中的一个电子并改变它的动量时，它也将散射电子对中的另一个电子，在相反方向引起动量的等量变化。因此，成对电子的平均运动不减慢也不加快，这就说明超导态的电子对运动时不消耗能量，因而表现出零电阻的特性，这也是超导体中可以产生永久电流的原因。

第 6 章　材料的介电性能

电介质是在电场作用下具有极化能力并在其中长期存在电场的一种物质，电介质具有极化能力和其中能够长期存在电场这种性质是电介质的基本属性，也是电介质多种实际应用的基础，电介质对电场的响应特性不同于金属导体。金属的特点是电子的共有化，体内有自由载流子，这样就决定了金属具有良好的导电性。然而，在电介质体内，一般情况下，只具有被束缚着的电荷，在电场的作用下将不能以传导方式而只是以感应的方式，即以其中正、负电荷受电场驱使形成正、负电荷中心不相重合的电极化方式来传递和记录电的影响。尽管对不同种类的电介质，电极化的机制各不相同，然而以电极化方式响应电场的作用，却是共同的。

电介质体内一般没有自由电荷，自然具有良好的绝缘性能。在工程应用上，常在需要将电路中具有不同电位的导体彼此隔开的地方使用电介质材料，就是利用介质的绝缘特性，从这个意义上讲，电介质又可称为绝缘材料或绝缘体。工程上实际应用的电介质毕竟与理想的电介质不同，最主要的是实际电介质在电场作用下存在泄露电流和电能的耗散以及在强电场下还可能导致电介质的破坏，因此，除要研究极化外，还要研究有关电介质的电导、损耗以及击穿特性，这些就是电介质物理需要研究的主要问题。通常采用电极化（介电常数）、电导、介电损耗和击穿 4 大常数来描述电介质的介电性能。

电介质种类繁多，在工程应用上，还常常按照组成电介质的分子的电荷在空间分布的情况进行分类。此分类方法，一般将电介质分为极性电介质和非极性（中性）电介质。当外电场作用时，介质由正、负电荷中心相重合的中性分子组成，这样的介质即为非极性（中性）介质，如聚四氟乙烯薄膜、变压器油等；若由正、负电荷中心不相重合的极性分子组成，这样的介质即为极性介质，如电容器纸的主要成分——纤维素以及聚氯乙烯薄膜等。

平行板电容器在工作时，它的两个金属板的相对两个表面上总是分别带上等量异号的电荷+q 和-q，这时两板间有一定的电势差 U（保持不变，形成静电场）。一个电容器所带的电量 q 与其电势差 U 之比为定值（电介质不变），这个比例定值称为电容器的电容（通常记做 C）。对真空平行板电容器

$$C_0 = \frac{A}{d}\varepsilon_0 \qquad (6\text{-}1)$$

式中，C_0 为真空平行板电容器的电容；A 为面积；d 为板极间距；ε_0 真空介电常数，$\varepsilon_0 = 8.85\times10^{-12}\mathrm{F\cdot m^{-1}}$。如果真空电容器中嵌入某种电介质，则

$$C = \frac{A}{d}\varepsilon = C_0 \frac{\varepsilon}{\varepsilon_0} = C_0 \varepsilon_r \tag{6-2}$$

式中，C 为嵌入电介质时电容器的电容；ε 为电介质的介电常数；ε_r 为电介质的相对介电常数。ε_r 反映了电介质极化的能力。

6.1　电介质的极化

6.1.1　极化现象及物理量

1. 极化现象

电介质最重要的性质就是在外电场作用下能够极化。所谓的极化从微观上说，一方面就是非极性质点（原子、分子、离子）正负电荷重心的分离，从而转变成偶极子。在电场作用下，构成质点的正负电荷在电场方向的有限范围内短程移动，组成一个偶极子。设正电荷与负电荷在外电场方向上的位移矢量为 $\boldsymbol{l}$（方向由负电荷指向正电荷），则定义此偶极子的电偶极矩（感应极矩）$\boldsymbol{\mu}= q\boldsymbol{l}$，规定其方向从负电荷指向正电荷，即电偶极矩的方向与外电场 $\boldsymbol{E}$ 的方向一致。另一方面，如果介质中含有极性分子，则这些极性分子都可看作偶极子。在外电场作用下，极性分子发生转向，转向的结果是每一个极性轴趋于电场方向，每一个偶极子的电偶极矩应看作原极性分子偶极矩在电场方向的投影。而感生电偶极矩则是极性分子在加外电场后偶极矩在电场方向投影的增加量。

2. 极化相关的物理量（在静电场中）

单位电场强度下，质点电偶极矩（指感应极矩，下同）的大小称为质点的极化率，即

$$\alpha = \frac{\mu}{E_{\text{loc}}} \tag{6-3}$$

式中，α 为质点的极化率，表征材料的极化能力，只与材料的性质有关，其单位为 F·m^2；μ 为质点电偶极矩大小，其单位为 C·m；E_{loc} 为作用在微观质点上的局部电场，它与宏观外电场并不相同。

定义介质单位体积内的电偶极矩总和为介质的极化强度，即

$$\boldsymbol{P} = \frac{\sum \boldsymbol{\mu}}{V} \tag{6-4}$$

式中，$\boldsymbol{P}$ 为介质的极化强度，其单位为 C·m^{-2}；$\boldsymbol{\mu}$ 为质点的电偶极矩；V 为介质的体积；

如果介质单位体积中的极化质点数等于 n，每一偶极子的电偶极矩具有同一方向（电场方向），所表示矢量和的式（6-4）可用下面的标量式代替

$$P = n\mu = n\alpha E_{\text{loc}} \tag{6-5}$$

式中，μ 为各质点的平均偶极矩。

定义

$$P = \varepsilon_0 \chi E \tag{6-6}$$

式中，E 为宏观平均电场强度；χ 称为电介质极化系数，它将介质的宏观电场和宏观物理量极化强度联系起来。不同电介质有不同的极化系数 χ，可以证明极化系数 χ 和相对介电常数 ε_r 有如下关系

$$\chi = \varepsilon_r - 1 \tag{6-7}$$

电位移 $\boldsymbol{D}$ 是为了描述电介质的高斯定理所引入的物理量，其定义为

$$\boldsymbol{D} = \varepsilon_0 \boldsymbol{E} + \boldsymbol{P} \tag{6-8}$$

式中，$\boldsymbol{D}$ 为电位移；$\boldsymbol{E}$ 为电场强度；$\boldsymbol{P}$ 为电极化强度。式（6-8）描述了 $\boldsymbol{D}$、$\boldsymbol{E}$、$\boldsymbol{P}$ 三矢量之间的关系，这对于各向同性电介质和各向异性电介质都是适用的。对各向同性电介质，联系式（6-6）和式（6-8）可得

$$D = \varepsilon_0 E + P = \varepsilon_0 E + \varepsilon_0 \chi E = \varepsilon_0 \varepsilon_r E = \varepsilon E \tag{6-9}$$

如果是各向异性介质，如石英单晶体等，则 $\boldsymbol{P}$ 与 $\boldsymbol{D}$、$\boldsymbol{E}$ 的方向一般并不相同，极化系数 χ 也不能只用数值来表示，但式（6-8）仍适用。

6.1.2 宏观极化强度与微观极化率的关系（在静电场中）

1. 宏观电场 E

上面提到介质宏观平均电场强度的概念，对于介质内宏观电场的的贡献，一是外加电场（由物体外部固定的电荷在真空中所产生的电场）；二是构成物体的所有质点电荷的电场之和。

当介质极化后，它就有了极化强度 P，为了求出极化强度 P 对宏观电场的贡献，可以对样品所有偶极子的和加以简化，如图 6-1 所示。极化强度 P 造成的电场可认为是由表面束缚电荷引起的。根据静电学原理，由均匀极化所产生的电场等于分布在物体表面上的束缚电荷在真空中产生的电场。令此电场强度为 E_{d}，它和外加电场 E_0 方向相反，因而称之为退极化场，如图 6-1 所示。一个椭球形样品在外电场下能产生均匀的极化强度和均匀的退极化场。这样对宏观场的贡献完全来自外加电场 E_0 和退极化场强度 E_{d}（矢量和），即

$$E = E_0 + E_{\text{d}} \tag{6-10}$$

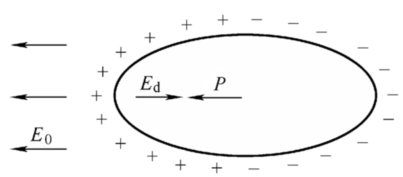


图 6-1 退极化场 E_{d}

2. 质点位置上的局部电场 E_{loc}

作用在一个质点位置上的局部电场（有时称为有效电场），其数值与宏观电场之间相差甚大。介质中作用于一个质点位置上的局部电场是外加电场 E_0 及介质中其他质点所

产生的电场之和。对其他质点的偶极子场求和的标准方法是以一个想象的参考质点为球心划出一个球，如图 6-2 所示。该圆球半径应比质点间距大得多，这样对讨论的球心质点来说，球外电介质可作为连续介质，即为均匀介质。同时，球半径又比整个介质（椭球样品）小得多，因此对宏观来说，可视球内为均匀的，即宏观电场对球内各点作用一样。一般可选球半径为质点间距的几十到几百倍。洛伦兹（H. A. Lorentz）设想把球挖空，使球外介质的作用归结为空球表面极化电荷作用场（E_1）和整个介质外边界表面极化电荷作用场（E_d）之和，球内则只考虑原点附近偶极子的影响即 E_2，如图 6-2 所示。所以对一个参考质点（球心）来说，局部电场为

$$E_{loc} = E_0 + E_d + E_1 + E_2 \tag{6-11}$$

式中，E_1 称为洛伦兹场。

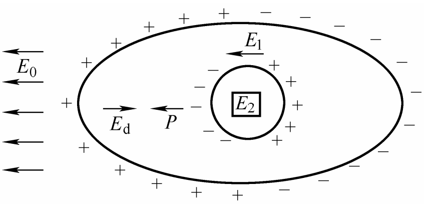


图 6-2　介质中质点上的内电场

洛伦兹曾计算此设想的空腔表面上的极化电荷所产生的电场 E_1。如图 6-3 所示，以 θ 表示相对于极化方向的夹角，θ 处空腔表面上的面电荷密度就是$-P\cos\theta$（P 为极化强度），取 dθ 角对应的微小环球面，其表面积为 dS。dS 面上的电荷 dq 为

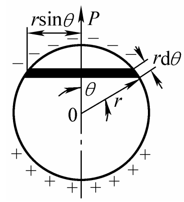


图 6-3　球形空腔电场计算

$$\mathrm{d}q = -P\cos\theta \mathrm{d}S = -2\pi r^2 P\cos\theta\sin\theta \mathrm{d}\theta \tag{6-12}$$

dq 在空腔球心 0 点产生的电场（在 P 方向上的投影）为

$$\mathrm{d}E = -\frac{1}{4\pi\varepsilon_0}\frac{\mathrm{d}q}{r^2}\cos\theta = \frac{1}{2\varepsilon_0}P\cos^2\theta\sin\theta \mathrm{d}\theta \tag{6-13}$$

则整个空腔球面上的电荷在球心产生的电场（洛伦兹场）为

$$E_1 = \int_0^{\pi}\mathrm{d}E = \int_0^{\pi}\frac{1}{2\varepsilon_0}P\cos^2\theta\sin\theta \mathrm{d}\theta = \frac{1}{3\varepsilon_0}P \tag{6-14}$$

现在来讨论球内其他质点所产生的电场 E_2。令球内其他质点（分子、原子、离子）j 的坐标为 x_j、y_j、z_j；而质点电矩 μ_j 的三分量为 μ_{jx}、μ_{jy}、μ_{jz}；则这些电矩在中心处所产生的电场沿外场方向（设为 x 向）的分量亦即所需求的 E_2，为

$$E_2=\frac{1}{4\pi\varepsilon_0}\sum_j\left(\mu_{jx}\frac{3x_j^{\ 2}-R_j^{\ 2}}{R_j^{\ 5}}+\mu_{jy}\frac{3x_jy_j}{R_j^{\ 5}}+\mu_{jz}\frac{3x_jz_j}{R_j^{\ 5}}\right) \tag{6-15}$$

式中，$R_j=(x_j^2+y_j^2+z_j^2)^{1/2}$，为电矩 μ_j 的位矢长度。对于由同种原子（或离子）组成的简立方晶体。设立方体的一边同宏观电场 E 的方向重合（在 x 向），则 $\mu_{jy}=\mu_{jz}=0$，令 a 为晶格常数，原子 j 的坐标为（al_{j1}，al_{j2}，al_{j3}），则

$$\begin{aligned}E_2&=\frac{1}{4\pi\varepsilon_0}\sum_j\mu_{jx}\frac{3a^2l_{j1}^{\ 2}-(l_{j1}^{\ 2}+l_{j2}^{\ 2}+l_{j3}^{\ 2})a^2}{(l_{j1}^{\ 2}+l_{j2}^{\ 2}+l_{j3}^{\ 2})^{5/2}a^5}\\&=\frac{1}{4\pi\varepsilon_0}\sum_j\frac{\mu_{jx}}{a^3}\frac{3l_{j1}^{\ 2}-(l_{j1}^{\ 2}+l_{j2}^{\ 2}+l_{j3}^{\ 2})}{(l_{j1}^{\ 2}+l_{j2}^{\ 2}+l_{j3}^{\ 2})^{5/2}}\end{aligned} \tag{6-16}$$

由于简单立方晶体的对称性，在球壳内，有

$$\sum_j l_{j1}^{\ 2}=\sum_j l_{j2}^{\ 2}=\sum_j l_{j3}^{\ 2}=\frac{1}{3}\sum_j(l_{j1}^{\ 2}+l_{j2}^{\ 2}+l_{j3}^{\ 2}) \tag{6-17}$$

故式（6-16）等于零，即

$$E_2=0 \tag{6-18}$$

同样，对于体心立方、面心立方，以及氯化钠结构的晶体，也可以证明 $E_2=0$，对于具有这类结构的晶体，有效场

$$E_{\text{loc}}=E_0+E_{\text{d}}+\frac{1}{3\varepsilon_0}P=E+\frac{1}{3\varepsilon_0}P \tag{6-19}$$

式中，E 为宏观电场。这就是洛伦兹关系。关于立方对称离子晶体的实验数据证明洛伦兹关系是正确的。

决不能认为所有具有立方结构的晶体，它们的 E_2 都等于零。例如钛酸钡（$BaTiO_3$）虽属立方结构，但其中氧的周围并不是立方对称的，因而 E_2 不等于零。所以在应用式（6-19）时必须慎重。事实上，对于具体的问题应该从式（6-15）出发实际计算 E_2。对于复式格子，不同原子（或离子）的周围情况不同，作用在不等同的原子（或离子）处的有效场不同，对于 E_2 应该就不同中心原子或离子进行计算。

另外必须指出的是，上面的考虑近似地适合于电子或离子的位移极化，而对于固有电矩的转向极化，上述的有效场则全然失效，这是因为洛伦兹模型中在考虑球外对球内中心分子电矩的长程作用时，过多地计算了转向的影响。

3. 克劳修斯—莫索蒂方程

根据 D、E 和 P 的关系可知

$$P=D-\varepsilon_0E=(\varepsilon-\varepsilon_0)E=\varepsilon_0(\varepsilon_r-1)E \tag{6-20}$$

将式（6-20）代入式（6-19）得

$$E_{\mathrm{loc}} = E + \frac{1}{3\varepsilon_0}\varepsilon_0(\varepsilon_r - 1)E = \frac{\varepsilon_r + 2}{3}E \tag{6-21}$$

由式（6-5）、式（6-20）和式（6-21）可得

$$P = n\alpha E_{\mathrm{loc}} = n\alpha \times \frac{\varepsilon_r + 2}{3}E = \varepsilon_0(\varepsilon_r - 1)E \tag{6-22}$$

即

$$\frac{\varepsilon_r - 1}{\varepsilon_r + 2} = \frac{n\alpha}{3\varepsilon_0} \tag{6-23}$$

此式称为克劳修斯—莫索蒂方程。它建立了宏观量 ε_r 与微观量 α 之间的关系。此式适用于分子间作用很弱的气体、非极性液体和非极性固体以及一些型离子晶体和具有适当对称的晶体。

对具有两种以上极化质点的介质，上式可变为

$$\frac{\varepsilon_r - 1}{\varepsilon_r + 2} = \frac{1}{3\varepsilon_0}\sum_k n_k\alpha_k \tag{6-24}$$

式中，n_k、α_k 分别为介质中的第 k 种极化质点的单位体积数量和质点的极化率。由式（6-20）可看出，为了获得高介电常数，除了选择 α 大的离子外，还要求 n 大，即单位体积的极化质点数要多。

6.1.3　电介质极化机制

电介质在外加电场作用下产生宏观的电极化强度，实际上是电介质微观上各种极化机制贡献的结果，它包括电子的极化、离子的极化（两者均又可分为位移极化和松弛极化）、电偶极子取向极化和空间电荷极化等。位移极化是一种弹性的、瞬时完成的极化，不消耗能量。松弛极化与热运动有关，完成这种极化需要一定的时间，并且是非弹性的，因而消耗一定的能量。

1. 电子、离子位移极化

（1）电子位移极化

在外电场作用下，原子外围的电子云相对于原子核发生位移，原子中的正、负电荷重心产生相对位移，这种极化称为电子位移极化。电子位移极化的性质具有一个弹性束缚电荷在强迫振动中所表现出来的特性。因为电子很轻，它们对电场的反应很快，可以光频跟随外电场变化。根据玻尔原子模型的经典理论可以计算出电子的平均极化率 α_{e}。

一个点电荷（电子）$-e$ 环绕以点电荷（原子核）$+q$ 为圆心的圆周轨道运行。垂直于轨道平面的电场 E_{loc} 使 $-e$ 的轨道中心沿轴线从轨道中心 M 移至点 O，移动距离为 d，如图 6-4 所示。则原子感生偶极矩为 $\mu = ed$。由图 6-4 可见，$d/R = eE_{\mathrm{loc}}/F_R$，$F_R$ 为沿轨道运行电子的离心力。由于 $R \gg d$，则 $F_R \approx e^2/(4\pi\varepsilon_0 R^2)$，因此有

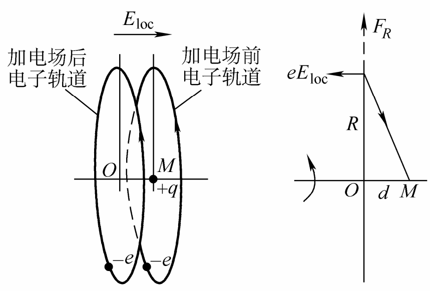


图 6-4　轨道模型

$$\frac{d}{R}=\frac{eE_{loc}}{F_R}=\frac{4\pi\varepsilon_0 R^2 E_{loc}}{e} \tag{6-25}$$

则

$$\mu = ed = 4\pi\varepsilon_0 R^3 E_{loc} \tag{6-26}$$

电子的极化率 α_e 为

$$\alpha_e = 4\pi\varepsilon_0 R^3 \tag{6-27}$$

可见，电子极化率的大小与原子（离子）的半径有关。

若考虑同类原子的一个集合，它们所有轨道是随机取向，则在电场方向上平均感生偶极矩为 $<\mu>=\mu<\cos^2\theta>$，尖括号表示对其中的物理量求平均值，θ 为偶极矩与电场方向的夹角。如电场强度较低，则 $<\cos^2\theta>=1/3$，所以

$$<\mu>=\frac{4}{3}\pi\varepsilon_0 R^3 E_{loc} \tag{6-28}$$

则

$$\alpha_e = \frac{4}{3}\pi\varepsilon_0 R^3 \tag{6-29}$$

对 O^-、F^-、Na^+、Mg^{2+}、S^-、Cl^-、K^+ 和 Ca^{2+} 离子半径的 X 射线测量值与经典理论计算值比较表明，除 Mg^{2+}、Ca^{2+} 离子外，相符较好，这说明经典模型是有效的。

还可以从另外一个角度估算电子的极化率 α_e 的大小，如导出电子的极化率 α_e 和电子与离子组成的弹性偶极子的固有频率的关系。设电子质量为 m，电子轨道中心在垂直于轨道平面的交变电场 $E_{loc}=E_0 e^{i\omega t}$（这里 E_0 为交变电场的幅值，ω 为电场交变频率，t 为时间）作用下位移为 x。电子轨道中心与原子核（离子）的弹性恢复系数为 β。则电子的运动方程为

$$m\frac{d^2x}{dt^2} = -\beta x = -eE_0 e^{i\omega t} \tag{6-30}$$

这个振动方程的解为

$$x = \left(\frac{-e}{\beta - m\omega^2}\right) E_0 \mathrm{e}^{\mathrm{i}\omega t} \tag{6-31}$$

由式（6-31）得电偶极矩为

$$\mu = -ex = \frac{e^2}{m}\left(\frac{1}{\beta / m - \omega^2}\right) E_0 \mathrm{e}^{\mathrm{i}\omega t} = \frac{e^2}{m}\left(\frac{1}{\omega_0{}^2 - \omega^2}\right) E_{\mathrm{loc}} \tag{6-32}$$

式中，$\omega_0 = \sqrt{\beta / m}$，即为电子与离子组成的弹性偶极子的固有频率。由式（6-32）可得电子的极化率 α_{e} 为

$$\alpha_{\mathrm{e}} = \frac{e^2}{m}\left(\frac{1}{\omega_0{}^2 - \omega^2}\right) \tag{6-33}$$

令 $\omega \to 0$，得静态极化率 $\alpha_{\mathrm{e}0}$ 为

$$\alpha_{\mathrm{e}0} = \frac{e^2}{m\omega_0{}^2} \tag{6-34}$$

由式（6-34）可以看出电子极化率依赖于频率。式（6-33）中极化率与频率的关系反映了极化的惯性。测量电子极化一般在光频（紫光）下进行，此时，其他极化机构（分子、离子极化）由于惯性跟不上电场的变化，因而此时的介电常数 ε_∞（∞ 表示交变电场频率很大）几乎完全来自电子极化率的贡献。

（2）离子位移极化

在离子晶体和玻璃等无机电介质中，正、负离子处于平衡状态，其偶极矩的矢量和为零。但这些离子，在电场作用下，除了离子内部产生电子位移极化外，离子本身将发生可逆的弹性位移。正离子沿电场方向移动，负离子沿反电场方向移动，正、负离子发生相对位移，形成感应偶极矩，这就是离子位移极化。离子的相对位移包括离子之间的间距变化和键与键之间夹角的变化。

为简单起见，取由两个异性离子（正负离子的质量分别为 M_+和 M_-，所带电量绝对值均为 q）组成的分子来研究离子位移极化率 α_{i}。在平行于正、负离子连线方向上施加电场 $E_{\mathrm{loc}} = E_0\mathrm{e}^{\mathrm{i}\omega t}$，与电子位移极化类似，在电场中离子的位移，仍然受到弹性恢复力的限制（弹性恢复系数为 β）。设正离子位移 δ_+，负离子位移 δ_-（δ_+与外电场方向相同为正，δ_-与外电场方向相反为负），如图 6-5 所示。则感生的电偶极矩 μ_{i} 为

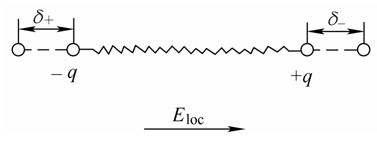


图 6-5　离子位移极化模型

$$\mu_{\mathrm{i}} = q(\delta_+ - \delta_-) = \alpha_{\mathrm{i}} E_{\mathrm{loc}} \tag{6-35}$$

在交变电场作用下，两离子的运动方程可写为

$$\left.\begin{aligned} M_+\frac{\mathrm{d}^2\delta_+}{\mathrm{d}t^2} &= -\beta(\delta_+-\delta_-)+qE_0\mathrm{e}^{\mathrm{i}\omega t} \\ M_-\frac{\mathrm{d}^2\delta_-}{\mathrm{d}t^2} &= -\beta(\delta_--\delta_+)-qE_0\mathrm{e}^{\mathrm{i}\omega t} \end{aligned}\right\} \tag{6-36}$$

引入相对运动约化质量 M*为

$$M^* = \frac{M_+M_-}{M_+ + M_-} \tag{6-37}$$

则由式（6-36）得

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}(\delta_+-\delta_-) = -\frac{\beta}{M^*}(\delta_+-\delta_-)+\frac{q}{M^*}E_0\mathrm{e}^{\mathrm{i}\omega t} \tag{6-38}$$

设 $\omega_0 = \sqrt{\beta / M^*}$ 为相对振动的固有频率，则式（6-38）的解为

$$(\delta_+-\delta_-) = \frac{q}{M^*}\left(\frac{1}{{\omega_0}^2-\omega^2}\right)E_0\mathrm{e}^{\mathrm{i}\omega t} \tag{6-39}$$

则离子极化率 α_i 为

$$\alpha_\mathrm{i} = \frac{q^2}{M^*}\left(\frac{1}{{\omega_0}^2-\omega^2}\right) \tag{6-40}$$

令 $\omega\to0$，得静态极化率

$$\alpha_{\mathrm{i}0} = \frac{q^2}{M^*{\omega_0}^2} = \frac{q^2}{\beta} \tag{6-41}$$

可见，离子位移极化和电子位移极化的表达式一样，都具有弹性偶极子的极化性质。ω_0 可由晶格振动红外吸收频率测量出来，从而得到离子位移极化建立的时间约为 10^{-12}～10^{-13}s。

2. 电子、离子和极性分子松弛极化

这种极化机制也是由外加电场造成的，但与带电质点的热运动状态密切相关。例如，当材料中存在着弱联系的电子、离子和偶极子等弛豫质点时，温度造成的热运动使这些质点分布混乱，而电场使它们有序分布，平衡时建立了极化状态。这种极化具有统计性质，称为热松弛（弛豫）极化。极化造成带电质点的运动距离可与分子大小相比拟，甚至更大。

由于是一种弛豫过程，建立平衡极化时间约为 10^{-2}～10^{-9}s，并且由于创建平衡要克服一定的位垒，故吸收一定能量，因此，与位移极化不同，松弛极化是一种非可逆过程。

松弛极化包括电子松弛极化、离子松弛极化、偶极子松弛极化。多发生在聚合物分子、晶体缺陷区或玻璃体内。

（1）电子松弛极化

由于晶格的热振动、晶格缺陷、杂质引入、化学成分局部改变等因素，使电子能态发生改变，出现位于禁带中的局部能级形成所谓弱束缚电子。例如色心点缺陷之一的“F—心”就是由一个负离子空位俘获了一个电子所形成的。“F—心”的弱束缚电子为周围结点上的阳离子所共有，在晶格热振动下，可以吸收一定能量由较低的局部能级跃迁到较高的能级而处于激发态，连续地由一个阳离子结点转移到另一个阳离子结点，类似于弱联系离子的迁移。外加电场使弱束缚电子的运动具有方向性，这就形成了极化状态，称之为电子松弛极化。

电子松弛极化与电子位移极化不同，是一种不可逆过程。由于这些电子处于弱束缚状态，因此，吸收一定的能量，电子可作短距离运动，运动距离比位移极化的电子大得多，所以极化作用也强烈得多。

但弱束缚电子和自由电子不同，不能远程迁移。因此电子松弛极化和电导不同，只有当弱束缚电子获得更高的能量时，受激发跃迁到导带成为自由电子，才形成电导。由此可见，具有电子松弛极化的介质往往具有电子电导特性。

电子松弛极化主要是折射率大、结构紧密、内电场大和电子电导大的电介质的特性。一般以 TiO_2 为基础的电容器陶瓷很容易出现弱束缚电子，形成电子松弛极化。含有 Nb^{+5}、Ca^{+2}、Ba^{+2} 杂质的钛质瓷和以铌、铋氧化物为基础的陶瓷，也具有电子松弛极化。

电子松弛极化建立的时间约为 10^{-2}～10^{-9}s，在电场频率高于 10^{9}Hz 时，这种极化就不存在了。

（2）离子松弛极化

和晶体中存在弱束缚电子类似，在晶体中也存在弱联系离子。在完整离子晶体中离子处于正常结点，能量最低最稳定，称之为强联系离子。它们在极化状态时，只能产生弹性位移，离子仍处于平衡位置附近。而在玻璃态物质、结构松散的离子晶体或晶体中的杂质或缺陷区域，离子自身能量较高，易于活化迁移，这些离子称弱联系离子。弱联系离子极化时，可以从一平衡位置移动到另一平衡位置。但当外电场去掉后离子不能回到原来的平衡位置，这种迁移是不可逆的，迁移的距离可达到晶格常数数量级，比离子位移极化时产生的弹性位移要大得多。然而需要注意的是弱离子弛豫极化不同于离子电导，因为后者迁移距离属远程运动，而前者运动距离是有限的，它只能在结构松散或缺陷区附近运动，越过势垒到新的平衡位置。其越过势垒 $U_{松}$要比电导势垒 $U_{电导}$小，如图 6-6 所示。

以离子晶体晶格缺陷为例，求解离子松弛极化率。如图 6-7 所示，设晶体中存在着正离子 M^{+}的空位（在 B 处），这种空位相当于一个负电荷。为了保持电中性，晶格中自然会出现电荷的某种补偿机构。例如，在空位周围出现 M^{++}。这样，M^{+}的空位和 M^{++}就构成电矩。由于 M^{+}的空位带负电，当施加外电场于晶体后，它

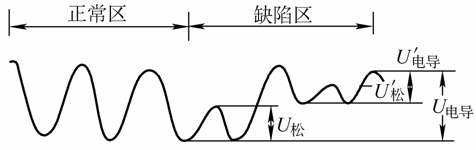


图 6-6　离子松弛极化与离子电导势垒

将逆外电场正的方向排列，如图 6-7 所示，在 A 位置上的离子 M^+易位到原来的空位 B 处，亦即正离子空位从位置 B 过渡到位置 A。这样，电矩可以基本上转到和外电场平行的方向。在趋于和外场平行的转动中，电矩必须克服势垒。由于热起伏，电矩克服势垒是可能的。由此可见，电矩由一平衡态过渡到另一平衡态之前必须等待若干时间；换言之，由于电矩在两个或两个以上的平衡位置转动而引起的极化是一个弛豫过程。

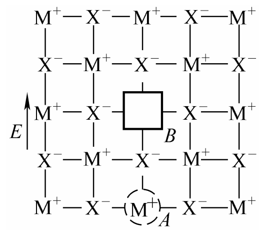


图 6-7　空位的移位导致电矩转向

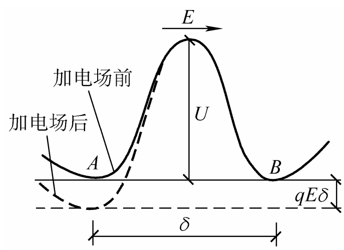


图 6-8　施加外电场后势垒变化

设 A、B 间的势垒为 U，如图 6-8 所示。以 v_0'（$=\omega_0'/(2\pi)$，ω_0'为角频率）代表离子的振动频率，沿某一方向运动的频率为 $v_0 = v_0'/6$。那么，在未加外电场时电矩（即正离子空位）从 A 向 B（或从 B 向 A）过渡的几率 $P_{AB(0)}$（$P_{BA(0)}$）为

$$P_{AB(0)} = P_{BA(0)} = v_0 \exp\left(-\frac{U}{kT}\right) \tag{6-42}$$

式中，T 为绝对温度；k 为玻耳兹曼常数。施加外电场 E（局部电场）后，A 处的势能比 B 处降低了 $qE\delta$（q 为离子空位电荷，δ 为 A、B 两点的距离），如图 6-8 所示。电矩由 A 向 B 过渡的几率变为 P_{AB} 为

$$P_{AB} = v_0 \exp\left(-\frac{U+qE\delta}{kT}\right) \tag{6-43}$$

加电场后，由 B 向 A 的过渡几率 P_{BA} 与未加电场的相同，即 $P_{BA} = P_{BA(0)}$，因而，电矩处位置 A 的几率就增加了。如果 $qE\delta \ll kT$，则

$$P_{AB} = v_0 \exp\left(-\frac{U+qE\delta}{kT}\right) \approx v_0\left(1-\frac{qE\delta}{kT}\right)\exp\left(-\frac{U}{kT}\right) \tag{6-44}$$

今设单位体积有 n_A 个正离子空位处于 A 位，而有 n_B 个正离子空位处于 B 位，则

$$\left.\begin{aligned}\frac{\mathrm{d}n_A}{\mathrm{d}t} &= -n_A P_{AB} + n_B P_{BA}\\ \frac{\mathrm{d}n_B}{\mathrm{d}t} &= +n_A P_{AB} - n_B P_{BA}\end{aligned}\right\} \tag{6-45}$$

设 n 为正离子总数，则 $n=n_A+n_B$。由式（6-45）得

$$\begin{aligned}\frac{\mathrm{d}(n_A-n_B)}{\mathrm{d}t}&=2n_BP_{BA}-2n_AP_{AB}\\&=(n_A-n_B)(P_{BA}+P_{AB})+n(P_{BA}-P_{AB})\end{aligned} \tag{6-46}$$

由式（6-42）、式（6-44）及 $qE\delta \ll kT$ 可得

$$P_{AB}+P_{BA}=\left(1-\frac{qE\delta}{kT}\right)P_{BA}+P_{BA}\approx 2P_{BA} \tag{6-47}$$

$$P_{BA}-P_{AB}=P_{BA}-\left(1-\frac{qE\delta}{kT}\right)P_{BA}=\frac{qE\delta}{kT}P_{BA} \tag{6-48}$$

由式（6-47）及式（6-48），式（6-46）变为

$$\frac{\mathrm{d}(n_A-n_B)}{\mathrm{d}t}=2(n_A-n_B)P_{BA}+nP_{BA}\frac{qE\delta}{kT} \tag{6-49}$$

设时间 $t=0$ 时，$n_A=n_B=n/2$，且设 $t=0$ 时加上外电场。在此条件下，式（6-49）的解为

$$n_A-n_B=\frac{nqE\delta}{2kT}\left[1-\exp(-2P_{BA}t)\right] \tag{6-50}$$

令 $\tau=\frac{1}{2\nu_0}\exp\left(\frac{U}{kT}\right)$（$\tau$ 为弱联系离子松弛（弛豫）时间常数），那么 $\tau=1/(2P_{BA})$，式（6-50）变为

$$n_A-n_B=\frac{nqE\delta}{2kT}\left[1-\exp\left(-\frac{t}{\tau}\right)\right] \tag{6-51}$$

$\Delta n=n_A-n_B$ 实际上为逆电场方向转移的净离子空位数。从式（6-51）可以看出，$t\to\infty$ 时，Δn 才稳定。实际上，$t=3\tau$ 时，极化就基本完成了。极化强度 P 和热松弛极化率 α_T 分别为

$$P_{t\to\infty}=\Delta n_{t\to\infty}q\delta=\frac{nq^2\delta^2}{2kT}E \tag{6-52}$$

$$\alpha_T=\frac{nq^2\delta^2}{2kT} \tag{6-53}$$

温度越高，热运动对质点的规则运动阻碍越强，因而 α_T 减小。

由计算可知，离子松弛极化率比电子位移极化率以及离子位移极化率大一个数量级，因而导致较大的介电常数。

在松弛极化 P 和介电常数 ε 与温度的关系中往往出现极大值。这是因为，一方面，温度升高，τ 减小，松弛过程加快，极化建立得更充分些，这时 ε 可升高，

另一方面，温度升高，极化率 α_T 下降，使 ε 降低，所以在适当温度下，ε 有极大值。一些具有离子松弛极化的陶瓷材料，其 ε—T 关系中未出现极大值，这是因为参加松弛极化的离子数随温度升高连续地增加。

离子松弛极化随频率的变化，在无线电频率下比较明显。由于一般松弛时间长达 10^{-2}～10^{-5}s，所以在无线电频率下（10^6Hz）离子松弛极化来不及建立，因而介电常数随频率升高明显下降。频率很高时，无松弛极化，只存在电子和离子位移极化（ε 趋近于 ε_∞）。

（3）极性分子固有电矩转向松弛极化

分子具有固有电矩，而在外电场作用下，电矩的转向所产生的电极化称为转向极化。许多电介质，例如一些有极的液体，具有较大的介电常数，这是与其中存在固有电矩有关的。

如果分子具有固有电矩，则在外电场作用下，除贡献电子极化和离子极化外，它们的固有电矩将趋于转到与外电场平行的方向，使介质的极化强度增大，如图 6-9 所示。特别重要的是，由于固有电矩间的相互作用具有长程有序的性质，一个分子的转向会带来周围许多分子的转向，这样会使得介电常数具有较大的数值。

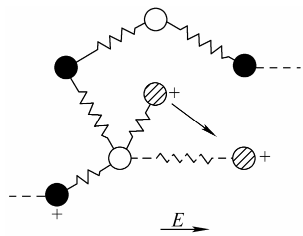


图 6-9　极性分子的转向极化

分子中固有电矩的存在，是由于分子结构的不对称性。例如水和氯化氢都有固有电矩。前者是由于 H_2O 分子并非直线结构（两个 OH 键夹角为 105°），后者则由于电荷在原子周围分布的不均衡，在氢的一端具有较多的正电荷所致。又如 CH_3CH_2Br 中，由于溴取代氢原子的结果，使得原来对称的分子 CH_3CH_3 变为不对称，因而也具有固有电矩。

下面来考虑固有电矩在外电场作用下的转向，从而求出其极化率 α。在这里的初步考虑中，将忽略固有电矩的相互作用（实际上一个分子的转向会带来周围许多分子转向），这只适用于稀疏情况下的气体。设气体包含大量相同的分子，而每个分子的固有电矩为 μ，在没有外电场作用时，由于热运动，这些电偶极子的排列是完全无规则的，因而就整个气体来看，并不具有电矩。当加上外电场 E（局部电场）后，每个电矩都受到电场的作用，趋于有序化。另一方面，热运动使电矩无序化。可见同时存在有序化和无序化相矛盾的两个方面。在一定的温度 T 和一定的电场 E 下，两方面的作用达到暂时的互相平衡。

固有电矩 μ 在外电场 E 的热能为

$$u = -\boldsymbol{\mu}\boldsymbol{E} = -\mu E\cos\theta \tag{6-54}$$

式中，θ 为 $\boldsymbol{\mu}$ 与 $\boldsymbol{E}$ 间的夹角。根据玻耳兹曼统计，$\boldsymbol{\mu}$ 与 $\boldsymbol{E}$ 间的夹角介于 θ 和（θ+dθ）之间的几率正比于 $2\pi\delta\theta\mathrm{d}\theta\exp[\mu E\cos\theta/(kT)]$，因此，沿电场方向的平均电矩为

$$\mu < \cos\theta > = \frac{\int_0^{\pi} \mu \cos\theta \sin\theta \exp\left(\frac{\mu E \cos\theta}{kT}\right) \mathrm{d}\theta}{\int_0^{\pi} \mu \sin\theta \sin\theta \exp\left(\frac{\mu E \cos\theta}{kT}\right) \mathrm{d}\theta} \tag{6-55}$$

对很大电场 E，存在 $\mu E \gg kT$，有 $<\cos\theta> \approx 1$，这时电场有序化运动超过热无序化，使得所有的电矩都完全平行于电场方向，因此 $\mu<\cos\theta> = \mu$。

在 E 不太大，T 不太小时，热运动的无序化作用占相当优势的情况下，可以认为 $\mu E \ll kT$，有 $<\cos\theta> = \frac{\mu E}{3kT}$。最终可得固有电矩的转向极化率 $\alpha = \frac{\mu^2}{3kT}$。

转向极化一般需要较长时间，约为 10^{-2}～10^{-10}s。对于一个典型的偶极子，$\mu = e \times 10^{-10}$C·m，因此 $\alpha \approx 2 \times 10^{-38}$F·m^2，比电子极化率（$10^{-40}$F·m^2）高。

分子的总极化率可以认为是电子极化率、离子极化率和转向极化率的总和。对于单位体积分子数为 n 的体系，存在关系 $\frac{\varepsilon_r - 1}{\varepsilon_r + 2} = \frac{n\alpha}{3\varepsilon_0}$。各种极化机制随外电场变化的速度不同，频率较低时，三种极化机制均起作用，随外电场频率增大，固有电偶极矩的取向迟缓而不能跟上电场的变化，频率再高，离子的位移也不能跟上电场的变化，此时电子位移极化起主要作用。

3. 空间电荷极化

空间电荷极化常常发生在不均匀介质中。在电场作用下，不均匀介质内部的正负间隙离子分别向负、正极移动，引起介质内各点离子密度变化，即出现电偶极矩。这种极化叫做空间电荷极化。在电极附近积聚的离子电荷就是空间电荷。

实际上晶界、相界、晶格畸变、杂质等缺陷区都可成为自由电荷（间隙离子、空位、引入的电子等）运动的障碍。在这些障碍处，自由电荷积聚，形成空间电荷极化。宏观不均匀性，例如夹层、气泡，也可形成空间电荷极化。所以上述极化又称界面极化。由于空间电荷的积聚，可形成很高的与外电场方向相反的电场，因此这种极化有时称为高压式极化。

空间电荷极化随温度升高而下降。因为温度升高，离子运动加剧，离子扩散容易，因而空间电荷减小。

空间电荷的建立需要较长的时间，大约几秒到数十分钟，甚至数十小时，因而空间电荷极化只对直流和低频下的介电性质有影响。

4. 自发极化和高介晶体极化

以上介绍的各种极化机构是介质在外电场作用下引起的，没有外加电场时，这些介质的极化强度等于零。还有一种极化叫自发极化，这是一种特殊的极化形式。这种极化状态并非由外电场引起，而是由晶体的内部结构造成的。在这类晶体中，每一个晶胞里存在有固有电矩。这类晶体称为极性晶体。自发极化现象通

常发生在一些特殊结构的晶体中。其机理将在铁电体部分讨论。

由实验得知，大部分离子晶体，例如碱卤晶体、碱土金属的氧化物和硫化物的相对介电常数 ε_∞约为 1.6～3.5，ε_r 约为 5～12。但是有少数晶体，如金红石（TiO_2）和钙钛矿（$CaTiO_3$）型晶体，其相对介电常数 ε_∞和 ε_r 都相当高，金红石多晶体的 $\varepsilon_\infty = 7.8$，$\varepsilon_r = 110$～114；钙钛矿晶体的 $\varepsilon_\infty = 5.3$，$\varepsilon_r = 150$。这类晶体的介电常数与温度的关系不大，即没有松弛极化的特性。为了说明这这类晶体介电常数大的原因，人们提出一个假设：这类晶体的晶体结构比较特殊，在外电场作用下，由于离子间的相互作用，引起了及其强烈的局部内电场。在此内电场的作用下，离子的电子壳层发生了强烈的变形，离子本身也发生了强烈的位移，这使材料具有很高的介电常数。

计算这类晶体的介电常数时，图 6-2 中球内原点附近偶极子的影响 E_2 不但不为零，而且有很大的数值。计算 E_2 时，可用式（6-16），但通常晶体中存在着好几种性质和相互位置不同的离子。为了研究方便，应把它们所建立的附加内电场区分开。Сканавц 导出了金红石晶体的 ε_r 为

$$\frac{\varepsilon_r - 1}{\varepsilon_r + 2} \approx \frac{n}{3\varepsilon_0} \frac{\alpha_1 + 2\alpha_2 + \alpha_i}{1 - \alpha_2 C_{22} - \alpha_i C_{21}} \tag{6-56}$$

式中，n 为单位体积内 TiO_2“分子”的个数，即 Ti^{4+}的个数；α_1、α_2 分别为 Ti^{4+}、O^{2-}的电子极化率；α_i 为 TiO_2“分子”在点阵中的离子位移极化率；C_{22}、C_{21} 为晶体的内电场结构系数。当离子位移极化率不存在时，即 $\alpha_i = 0$，则得到纯电子极化时的公式

$$\frac{\varepsilon_\infty - 1}{\varepsilon_\infty + 2} \approx \frac{n}{3\varepsilon_0} \frac{\alpha_1 + 2\alpha_2}{1 - \alpha_2 C_{22}} \tag{6-57}$$

比较式（6-56）和式（6-57）可见，只要加入不大的离子极化率 α_i，ε_r 比起 ε_∞来就剧增。Сканавц 根据式（6-56）计算的结果与实验结果很接近。

各种极化形式的综合比较如表 6-1 及图 6-10 所示。

表 6-1 各种极化形式的比较

极化形式	存在介质	发生频率范围	与温度的关系	能量消耗
电子位移极化	一切陶瓷介质	直流—光频	无关	无
电子松弛极化	钛质瓷、以高价金属氧化物为基的陶瓷	直流—超高频	随温度变化有极大值	有
离子位移极化	离子结构介质	直流—红外	温度升高极化增强	很微弱
离子松弛极化	离子结构玻璃、结构不紧密的晶体及陶瓷	直流—超高频	随温度变化有极大值	有
转向极化	有机材料	直流—超高频	随温度变化有极大值	有
空间电荷极化	结构不均匀的陶瓷	直流—高频	随温度升高而减弱	有
自发极化	温度低于居里点的铁电材料	直流—超高频	随温度变化有显著极大值	很大

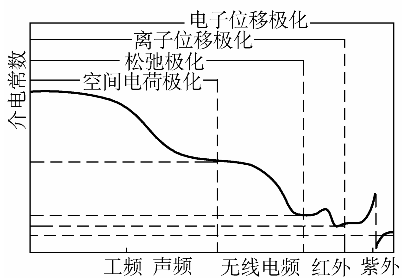


图 6-10　各种频率范围对介电常数的影响

6.2　交变电场下的电介质行为

本节将介绍在变动电场中电介质的行为。如电场不断地改变，介质内的极化也就要不断地改变。电场改变相当迅速时，极化就会追随不及而滞后。实际上介质中的多种极化是一些弛豫过程，从初态到末态要经历或多或少的弛豫时间。介质极化的这种弛豫，在变动电场中就引起介质损耗，并且使动态介电常数和静态介电常数不同。

6.2.1　复介电常数和介质损耗

对一个真空平行板电容器，加一交变电压 $U = U_0 e^{i\omega t}$（$U = U_0\cos\omega t$ 为交变电压的复数形式），则在电极上出现电荷 $Q = C_0U$，该电容上的电流 I_0 为

$$I_0 = \frac{dQ}{dt} = \frac{d(C_0U)}{dt} = i\omega C_0 U \tag{6-58}$$

式中，电容电流 I_0 超前电压 U 相位 90°，是一种非损耗性电流。

当电容器两极间充以非极性的完全绝缘的材料（电容为 C，介电常数为 ε，相对介电常数为 ε_r。对非极性的完全绝缘的材料，这些参数可视为不随外加电场频率变化。这些介质的极化完全是位移极化，介质内部未形成宏观的导电电流），加的电压不变，则电容电流 I_1 为

$$I_1 = \frac{dQ}{dt} = \frac{d(CU)}{dt} = i\omega CU \tag{6-59}$$

式中，$I_1 = \varepsilon_r I_0$，I_1 超前电压 U 相位 90°，是一种非损耗性电流。如果是非极性介质，内部存在自由电荷，设 G 是自由电荷的纯电导，则介质内存在电导电流分量 $I_2 = GU$。I_2 与 U 同相位，如图 6-11 所示。这时合成电流 I 为

$$I = I_1 + I_2 = (i\omega C + G)U \tag{6-60}$$

对平行板电容器，$G = \sigma A/d$，$C = \varepsilon A/d$，A 为板相对的面积，d 为板极间距。则总

电流密度 j 为

$$j=\frac{I}{A}=(\mathrm{i}\omega\varepsilon+\sigma)E \tag{6-61}$$

式中，E 为介质内宏观电场。于是可由 $j=\sigma^*E$ 定义复电导率 σ^*为

$$\sigma^*=\mathrm{i}\omega\varepsilon+\sigma \tag{6-62}$$

也可以由 $j=\mathrm{i}\omega\varepsilon^*E$ 定义复介电常数 ε^*为

$$\varepsilon^*=\frac{\sigma^*}{\mathrm{i}\omega}=\varepsilon-\mathrm{i}\frac{\sigma}{\omega} \tag{6-63}$$

损耗角 δ（图 6-11）由下式定义

$$\tan\delta=\frac{\text{损耗项}}{\text{电容项}}=\frac{\sigma}{\omega\varepsilon} \tag{6-64}$$

实际的介电材料是弱导电性的，或是极性的，或兼有此两种特性。加上交变电场后，电容器的电流可人为地分为三部分。一是介质的几何电容充电所造成的电流，即认为电介质为非极性的完全绝缘时的充电电流 I_1，方向垂直于 E（U）的方向（图 6-11）；二是由介质电导（漏导）造成的电流 I_2，与 E 同方向；三是由一些极化（弛豫极化、空间电荷极化等）的建立所造成的电流 I_3。由于这些极化有时跟不上电场的变化，造成 I_3 落后于 E 的相位，即与 E 有一夹角。那么 I_3 在 I_1 和 I_2 方向都有分量。因而，按上述方法导出实际介电材料的复数介电常数的实部不是精确等于上述的 ε，当然虚部也不是精确等于 σ/ω。

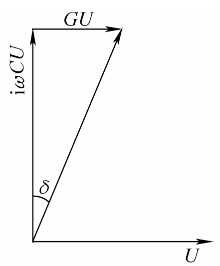


图 6-11 电容器上的电流

复介电常数普通的表达形式

$$\varepsilon^*=\varepsilon'-\mathrm{i}\varepsilon'' \tag{6-65}$$

这里，ε'和 ε''是依赖于频率的量。所以

$$\tan\delta=\frac{\varepsilon''}{\varepsilon'} \tag{6-66}$$

由此可知，损耗由复介电常数的虚部 ε''引起，通常电容电流由实部 ε'引起。

6.2.2 电介质弛豫和频率响应

前面介绍电介质极化微观机制时，曾分别指出不同极化方式建立并达到平衡时所需的时间。事实上只有位移极化可以认为是瞬时立即完成的（严格地讲电子位移极化是瞬时的），其他极化都需要时间，这样在交流电场作用下，电介质的极化就存在频率响应问题。

在一个实际介质的样品上突然加上一电场（阶跃电场），所产生的极化过程不是瞬时的，如图 6-12 所示。P_0 代表瞬时建立的极化（位移极化），P_1 代表松弛极化，P_1 是时间的函数（可写为 $P_1(t)$），它渐渐达到一稳定值 $P_{1\infty}$。这一滞后通常是由转向极化和空间电荷极化所致。在外电场施加和移去后，系统逐渐达到平衡状态的过程叫介质弛豫。由图 6-12 可以看出，t 时间极化强度 $P(t)$包括两项为

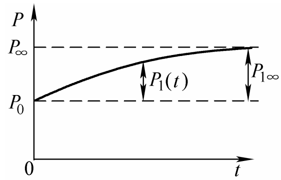


图 6-12　介质的弛豫过程

$$P(t) = P_0 + P_1(t) \tag{6-67}$$

当时间足够长时，$P_1(t) \to P_{1\infty}$，而 $P(t) \to P_\infty$。设 $P_0 = \varepsilon_0\chi_0 E$，$P_{1\infty} = \varepsilon_0\chi_1 E$（$E$ 为宏观电场强度，χ_0 为介质的瞬时极化系数，χ_1 为介质稳态极化系数）。根据弛豫过程的特征方程，有

$$\frac{\mathrm{d}P_1}{\mathrm{d}t} = \frac{P_{1\infty} - P_1}{\tau} \tag{6-68}$$

式中，τ 为弛豫时间常数。当 $t = 0$ 时，对于阶跃电场，有 $P_1(0) = 0$，因而可得

$$\begin{aligned} P(t) &= P_0 + P_1(t) = \varepsilon_0\chi_0 E + \varepsilon_0\chi_1[1 - \exp(-t/\tau)]E \\ &= \varepsilon_0\{\chi_0 + \chi_1[1 - \exp(-t/\tau)]\}E \end{aligned} \tag{6-69}$$

当外加电场是交变电场 $E = E_0 e^{i\omega t}$ 时，$P_1(t)$与 E 具有相同的变化频率（但相位角不一定相同），可设 $P_1(t)$为

$$P_1(t) = A\exp(i\omega t) \tag{6-70}$$

式中的 A 不一定为实数，$P_1(t)$与 E 的相位差可包括在 A 中。将式（6-70）代入式（6-68），并考虑到 $P_{1\infty} = \varepsilon_0\chi_1 E$，可得 A 为

$$A = \varepsilon_0\chi_1 E_0/(1 + i\omega\tau) \tag{6-71}$$

则

$$P_1(t) = \frac{\varepsilon_0\chi_1 E}{1 + i\omega\tau} \tag{6-72}$$

所以

$$P(t) = P_0 + P_1(t) = \varepsilon_0\left(\chi_0 + \frac{\chi_1}{1 + i\omega\tau}\right)E = \varepsilon_0\chi E \tag{6-73}$$

式中，χ 为介质的复极化系数。

则介质的复相对介电常数 ε_r 为

$$\varepsilon_r = 1 + \chi = 1 + \chi_0 + \frac{\chi_1}{1 + i\omega\tau} = \varepsilon_r' - i\varepsilon_r'' \tag{6-74}$$

式中，ε_r'和 ε_r''分别为复相对介电常数的实部和虚部。那么有

$$\left.\begin{aligned}\varepsilon_r' &= 1+\chi_0+\frac{\chi_1}{1+\omega^2\tau^2}\\ \varepsilon_r'' &= \frac{\chi_1\omega\tau}{1+\omega^2\tau^2}\end{aligned}\right\} \tag{6-75}$$

低频或静态时，$\varepsilon_r' = \varepsilon_{r0}$（$\varepsilon_{r0}$代表静态相对介电常数）；频率 $\omega\to\infty$时，$\varepsilon_r' = \varepsilon_{r\infty}$（$\varepsilon_{r\infty}$代表光频相对介电常数），所以有

$$\left.\begin{aligned}\varepsilon_{r0} &= 1+\chi_0+\chi_1\\ \varepsilon_{r\infty} &= 1+\chi_0\end{aligned}\right\} \tag{6-76}$$

由式（6-74）、式（6-75）、式（6-76），连同 $\tan\delta = \varepsilon_r'/\varepsilon_r''$，可得

$$\varepsilon_r = \varepsilon_{r\infty}+\frac{\varepsilon_{r0}-\varepsilon_{r\infty}}{1+\mathrm{i}\omega\tau} \tag{6-77}$$

$$\left.\begin{aligned}\varepsilon_r' &= \varepsilon_{r\infty}+\frac{\varepsilon_{r0}-\varepsilon_{r\infty}}{1+\omega^2\tau^2}\\ \varepsilon_r'' &= \frac{\varepsilon_{r0}-\varepsilon_{r\infty}}{1+\omega^2\tau^2}\omega\tau\\ \tan\delta &= \frac{\varepsilon_{r0}-\varepsilon_{r\infty}}{\varepsilon_{r0}+\varepsilon_{r\infty}\omega^2\tau^2}\omega\tau\end{aligned}\right\} \tag{6-78}$$

式（6-78）即为德拜公式。

式（6-78）可以分析描述电介质极化和频率、弛豫时间的关系。电介质的相对介电常数（实部和虚部）随所加电场的频率而变化。在低频时，相对介电常数大小与频率基本无关；当 $\omega\tau = 1$ 时，ε_r''有极大值；当 $\omega\tau = (\varepsilon_{r0}/\varepsilon_{r\infty})^{1/2}$ 时，$\tan\delta$ 有极大值。某种介质根据其 ε_{r0}、$\varepsilon_{r\infty}$、τ 值由式（6-78）作图，得到图 6-13 所示的三条曲线。

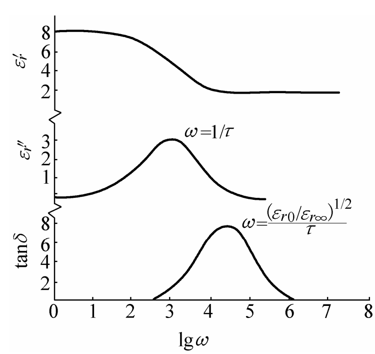


图 6-13 ε_r'、ε_r''、$\tan\delta$ 分别与 ω 的关系曲线

由于不同极化机制的弛豫时间不同，因此，在交变电场频率极高时，弛豫时间长的极化机制来不及响应所受电场的变化，故对总的极化强度没有贡献。图 6-14 为某一电介质的相对介电常数 ε_r（实部）、介电损耗 $\tan\delta$ 与频率 f 的关系。在图 6-14b 中，在偶极子松弛极化阶段、离子位移极化阶段（红外光谱范围）、电子位移极化阶段（紫外光谱范围）都有吸收峰。

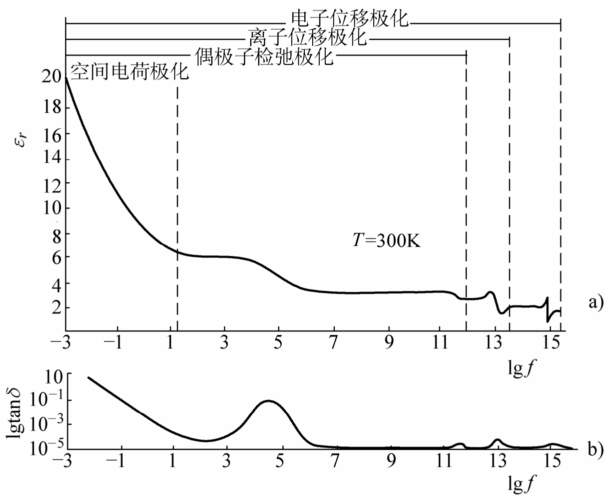


图 6-14　相对介电常数 ε_r（实部）与频率 f 关系曲线和介电损耗 $\tan\delta$ 与频率 f 关系曲线

6.2.3　介电损耗分析

1. 频率的影响

频率与介质损耗的关系，在德拜方程中有所体现。现分析如下。①当外加电场频率 ω 很低，即 $\omega\to 0$，介质的各种极化机制都能跟上电场的变化，此时不存在极化损耗，相对介电常数最大。介质损耗主要由电介质的漏电引起，则损耗功率 P_W 与频率无关；②当外加电场频率增加至某一值时（$\omega\tau < 1$），松弛极化跟不上电场变化，介电常数的实部减小，虚部增大，则 $\tan\delta$ 增大，P_W 也增大；③当频率 ω 提的很高，ε_r（实部）$\to\varepsilon_{r\infty}$，趋于最小值。由于此时 $\omega\tau\gg 1$，当 $\omega\to\infty$，则 $\tan\delta\to 0$；④当 ω 有适当值时，$\tan\delta$ 有最大值，如图 6-13 所示。此时，$\tan\delta$ 最大值主要由弛豫过程决定。如果介质电导显著增大，则 $\tan\delta$ 峰值变平坦，甚至没有最大值。

2. 温度的影响

温度对弛豫极化有影响，因此也影响到 P_W、ε_r 和 $\tan\delta$ 值的变化。温度升高，弛豫极化增加，而且离子间易发生移动，所以极化的弛豫时间 τ 减小，具体情况可结合德拜方程进行分析。①当温度很低时，τ 较大，由德拜方程可知，ε_r 较小，

$\tan\delta$ 较小。此时 $\omega^2\tau^2 \gg 1$，由式（6-78）可知，$\tan\delta \propto 1/(\omega\tau)$，$\varepsilon_r \propto 1/(\omega^2\tau^2)$。在这低温度范围内随温度上升，$\tau$ 减小，则 ε_r 和 $\tan\delta$ 上升，P_W 也上升；②当温度较高时，τ 较小，此时 $\omega^2\tau^2 \ll 1$，因此，随温度升高 τ 减小，$\tan\delta$ 减小。由于此时电导上升不明显，所以 P_W 主要决定于极化过程，P_W 也随温度上升而减小。联系低温部分可见，在某一温度 T_m 下，P_W 和 $\tan\delta$ 有极大值，如图 6-15 所示；③温度持续升高达很高时，离子热振动能很高，离子迁移受热振动阻碍增大，极化减弱，则 ε_r 下降，电导急剧上升，故 $\tan\delta$ 也增大（图 6-15）。

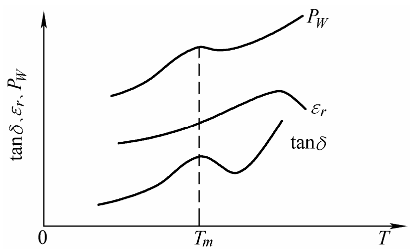


图 6-15　P_W、ε_r 和 $\tan\delta$ 与温度的关系曲线

从前面的分析可知，若电介质的电导很小，则弛豫极化损耗特征是，$\tan\delta$ 在与频率、温度的关系曲线中出现极大值。

3. 陶瓷材料的损耗

陶瓷材料的损耗主要来源于三部分：电导损耗；取向极化和弛豫极化损耗；电介质结构损耗。此外，无机材料表面气孔吸附水分，油污及灰尘等造成的表面电导也会引起较大的损耗。

概括地讲，极化损耗是指在极化过程中，带电质点（弱束缚电子和弱联系离子以及偶极子，或者空位等）移动时，由于与外电场作用不同步，而吸收了电场能量并把它传给周围的“分子”，使电磁能转变为“分子”的热振动能，把能量消耗在使电介质发热上。很明显，电导损耗是电流引起的材料电阻发热损耗。

而结构损耗主要是指陶瓷材料中玻璃相的损耗，以离子晶体为主晶相的陶瓷材料损耗主要来源于玻璃相。因此，高频瓷如氧化铝瓷、金红石瓷等都很少有玻璃相。

电工陶瓷中会有缺陷固溶体或多晶形转变，或者会有可变价离子如钛陶瓷等，往往具有显著的电子松弛极化损耗。

表 6-2 给出了一些电容器瓷的损耗角正切值。

表 6-2　电容器瓷的 $\tan\delta$（$f = 10^6$Hz，$T = 293\text{K} \pm 5\text{K}$）

材料	金红石瓷	钛酸钙瓷	钛酸锶瓷	钛酸镁瓷	钛酸锆瓷	锡酸钙瓷
$\tan\delta / \times 10^{-4}$	4～5	3～4	3	1.7～2.7	3～4	3～4

6.3　电介质在电场中的破坏

介质的特性，如绝缘、介电能力，都是指在一定的电场强度范围内的材料的特性，即介质只能在一定的电场强度以内保持这些性质。当电场强度超过某一临界值时，介质由介电状态变为导电状态，这种现象称为介电强度的破坏，或叫介质的击穿。相应的临界电场强度称为介电强度，或称为击穿电场强度。

对于凝聚态绝缘体，通常所观测到的击穿电场范围约为 10^5～$5\times10^6 V\cdot cm^{-1}$。从宏观尺度看，这些电场属于高电场，但从原子的尺度看，这些电场是非常低的。这清楚地表明，除了在非常特殊的实验室条件下，击穿决不是由于电场对原子或分子的直接作用所导致。电击穿是一种集体现象。能量通过其他粒子（例如，已经从电场中获得了足够能量的电子和离子）传送到被击穿的组分中的原子或分子上。

虽然严格地划分击穿类型是很困难的，但为了便于叙述和理解，通常将击穿类型分为三种：热击穿、电击穿和局部放电击穿。这里主要简述这几种击穿的机制。

6.3.1　热击穿机制

当固体电介质在电场作用下，由电导和介质损耗产生的热量超过试样通过传导、对流和辐射所能散发的热量时，试样中的热平衡就被破坏，试样温度不断上升，最终造成介质永久性的热破坏，这就是热击穿。显然，热击穿除与所加电压的大小、类型、频率和介质的电导、损耗有关外，还与材料的热传导、热辐射以及试样的形状、散热情况，周围媒质温度等等一系列因素有关。

为了定量地研究介质在什么条件下发生热击穿，即要获得热击穿判据，首先要建立介质在电场作用下的发热或散热的临界平衡方程，从而得到在各项条件下的热击穿电压。在电场作用下，如试样的发热功率为 W_1，散热功率为 W_2，临界热平衡方程为 $W_1(T_m) = W_2(T_m)$，其中，T_m 为介质达到临界热平衡时的最高极限温度，与 T_m 相应的电压就是热击穿电压。实际上要建立临界热平衡解析式是相当复杂的，现仅以一种最简单的情况为例加以说明。

设对面积为 A、厚度为 d 的试样施加直流电压 V，且介质只有漏导电流产生的热量。这时介质的电导率为 σ，则试样的电导 $G=\sigma A/d$，这时介质的发热功率 $W_1=GV^2$。介质的电导率 σ 随温度 T 的上升指数式增加，其关系式可表示为 $\sigma = A'\exp(-B/T)$，A'、B 为常数。则

$$W_1(V,T)=\frac{AA'}{d}V^2\exp\left(-\frac{B}{T}\right) \qquad (6\text{-}79)$$

上式表明，W_1 是电压 V 和介质温度 T 的函数，并随温度 T 增加指数上升。图 6-16 表示出了在不同的电压 V_1、V_2、V_3 下，W_1 与 T 关系的曲线。

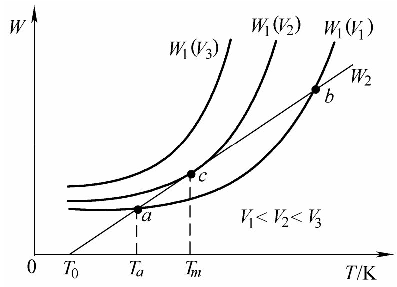


图 6-16　在电压作用下固体电介质的发热与散热曲线

介质中所产生的热量，一方面使试样本身的温度升高，另一方面通过热传导和热对流向周围散发热量。如环境温度为 T_0，散热功率 W_2 与温差 $T-T_0$ 成正比，则

$$W_2(T) = B'(T - T_0) \tag{6-80}$$

式中，B'为与热传导和热对流有关的散热系数，可以把 B'看成是一个与温度无关的常数。由上式可见，散热功率 W_2 与介质温度 T 呈线性，如图 6-16 所示。

由图 6-16 可见：①当外加电压 V_1 较小时，发热曲线 $W_1(V_1)$与散热直线 W_2 在 a 点相交。这表明，在 a 点发热功率 W_1 等于散热功率 W_2，介质处于热平衡情况，这时试样的温度为 T_a，电压可持续作用，温度不再升高；②若外加电压 V_3 较高，这时发热功率 W_1 恒大于散热功率 W_2，曲线 $W_1(V_3)$与直线 W_2 不可能相交，这表明介质在任何温度下都不会达到热平衡，这就使介质温度不断地升高，最后导致热击穿；③在一定电压 V_2 下，曲线 $W_1(V_2)$与直线 W_2 相切于 c 点，相应于 c 点介质的温度为 T_m，这就是介质达到临界热击穿时的最高极限温度。因为 $T<T_m$ 时，$W_1(V_2) > W_2$，温度继续上升至 T_m，在 T_m 维持热平衡，电压可持续作用，温度不再升高。而当 $T>T_m$ 时，$W_1(V_2) > W_2$，温度继续上升最后导致热击穿。发热曲线 $W_1(V_2)$是介质热稳定与不稳定状态的临界曲线，电压 V_2 就是固体电介质热击穿电压。由于热击穿电压不仅取决于固体电介质的本质，同时还取决于一系列外界因素，因此热击穿电压往往不作为表征介质特征的参数。

显然，当发热曲线 W_1 与散热直线 W_2 相切时，切点 c 应满足以下条件

$$\left.\begin{aligned} W_1(V,T)\Big|_{T=T_m} &= W_2(T)\Big|_{T=T_m} \\ \frac{\partial V_1}{\partial T}\bigg|_{T=T_m} &= \frac{\mathrm{d}V_2}{\mathrm{d}T}\bigg|_{T=T_m} \end{aligned}\right\} \tag{6-81}$$

这就是固体电介质的热击穿判据。

对于介质损耗较高的固体介质材料，在高频下的主要击穿形式是热击穿。

6.3.2　电击穿机制

在强电场下，固体导带中可能因冷发射或热发射存在一些电子。这些电子一方面在外电场作用下被加速，获得动能；另一方面与晶格振动相互作用，把电场能量传递给晶格。当两个过程在一定的温度和场强下平衡，固体介质有稳定的电导；当电子从电场中得到的能量大于传递给晶格振动的能量时，电子的动能就越来越大，至电子能量大到一定值时，电子与晶格振动的相互作用导致电离产生新电子，使自由电子数迅速增加，电导进入不稳定阶段，击穿发生。

1. 本征电击穿机制

实验上，本征电击穿表现的击穿主要是由所加电场决定的，在所使用的电场条件下，使电子温度达到击穿的临界水平。观察发现，本征击穿发生在室温或室温以下。发生的时间间隔很短，在微秒或微秒以下。本征击穿所以称之为“本征”，是因为这种击穿机制与样品或电极几何形状无关，或者与所加电场的波形无关。因此在给定温度下，产生本征击穿的电场值仅与材料有关。

这种击穿与介质中的自由电子有关。介质中自由电子的来源为杂质或缺陷能级、价带。

处理固体电介质击穿所使用的基本理论是基于方程式（击穿的极限条件）：$A(T_0, E, \alpha) = B(T_0, \alpha)$。式中，$A(T_0, E, \alpha)$为材料从所加电场获得的能量；$B(T_0, \alpha)$为材料消耗的能量；$T_0$为晶格温度；$E$为电场强度；$\alpha$为能量分布参数，决定于所采用处理的模型。

本征击穿理论可归结为基本上处理电子与晶格间能量的传递，并且考虑材料中电子能量的分布变化。参与能量传递作用的因素有：①偶极场中的晶格振动；②与偶极场晶格振动共有的电子壳层变形；③非偶极场短程电子轨道畸变。本征击穿理论所相关的电子能量分布变化的因素有：①电场对电子的加速作用；②传导电子间的碰撞；③传导电子与晶格的相互碰撞；④电子的电离、再复合和捕获；⑤电场梯度形成的扩散。

解决的问题是要求出能量平衡时临界电场强度，从而找出击穿时材料的临界温度。本征击穿机制有两种模型。一种是单电子近似模型。在这种模型中，单位时间单电子从电场中获得的能量为$A = e^2E^2\tau/m^*$。式中：e为电子电荷；m^*为电子有效质量；E为电场强度；τ为电子的平均弛豫时间。利用这种模型，说明材料在低温区，当温度升高时，引起晶格振动加强，电子散射增加，电子弛豫时间变短，因而使击穿电场反而提高。实验结果与之定性符合。这个模型仅适于材料本征击穿低温区。另一种是集合电子近似模型，考虑了电子之间的相互作用，建立了杂

质晶体电击穿的理论，根据这一模型导出

$$\ln E = c_0 \frac{\Delta u}{2kT_0} \tag{6-82}$$

式中，E 为击穿电场强度；k 为玻耳兹曼常数；T_0 为晶格温度；Δu 为能带中杂质能级激发态与导带底距离的一半。

热击穿也有与式（6-82）类似关系式，所以可以把本征击穿看成是热击穿的微观理论。

2. 雪崩式电击穿机制

热击穿机制对于许多陶瓷材料是适用的。如果材料尺寸可看成是薄膜时，则雪崩式击穿机制更为有效。

雪崩式电击穿机制是把本征电击穿机制和热击穿机制结合起来。因为当电子的分布不稳定时，必然产生热的结果。因此，这种理论是用本征电击穿理论描述电子行为，而击穿的判据采用的是热击穿性质。

雪崩式理论认为：电荷是逐渐或者相继积聚，而不是电导率的突然改变，尽管电荷集聚在很短时间内发生。

雪崩式电击穿最初的机制是场发射或离子碰撞。场发射假设由隧道效应来自价带的电子进入缺陷能级或进入导带，导致传导电子密度增加，其发射几率

$$P = aE\exp\left(-\frac{bI^2}{E}\right) \tag{6-83}$$

式中，E 为电场强度；a、b 为常数；I 为电流。由式（6-83）可见，只有当电场 E 相当强时，发射几率 P 才能高。采用脉冲热判据 $T=T_c$，T_c 是个临界参数，用式（6-83）估计出雪崩式击穿发生的临界场强为 $E = 10^7\text{V}\cdot\text{cm}^{-1}$。

Seitz 计算认为，只有达到每立方厘米 10^{12} 个电子的自由电子密度所具有的总能量，才能破坏电介质的晶格结构。当一个电子游离并开始碰撞，“解放”2 个电子，然后这 2 个依次去“解放”4 个，等等，只有进行 40 代这种碰撞才能导致“雪崩式”的击穿。这种简单的处理，使我们得知，临界电场依赖于样品厚度。样品至少有足够的厚度，达到 40 倍电子的平均自由程。也就是说，由阴极出发的初始电子在向其阳极运动的过程中，1cm 内的电离次数达 40 次，介质便击穿了。Seitz 估计粗糙一点，但可用来说明“雪崩”击穿的形成，并称之为“四十代理论”。

由“四十代理论”可以推论，当介质很薄时，碰撞电离不足以发展到 40 代，电子雪崩系列已进入阳极复合，此时介质不能击穿，因为这时电场强度不够高，因此便定性解释了薄层介质具有较高的击穿电场的原因。

雪崩电击穿和本征电击穿一般难于区分，在理论上它们的关系十分明显：本征电击穿理论中增加传导电子是继稳态破坏后突然发生的，而“雪崩”击穿是考虑高场强时，传导电子倍增过程逐渐达到难以忍受的程度，导致介质晶格破坏。

6.3.3　局部放电击穿

局部放电就是在电场作用下，在电介质局部区域中所发生的放电现象，这种放电没有电极之间形成贯穿的通道，整个试样并没有被击穿。例如气体的电晕放电、液体中的气泡放电都是局部放电。对于固体电介质来说，电极与介质之间常常存在着一层环境媒质：气隙或油膜。就固体电介质本身来说，实际上也是不均匀的，往往存在着气泡、液珠或其他杂质和不均匀的组分等。例如陶瓷就是一种多孔性的不均匀材料。由于气体和液体介电常数较小，因此承受的电场强度较高。同时气体和液体的击穿电场强度又比较低，于是当外施电压达到一定数值时，在这薄弱的区域，就发生局部放电。

局部放电是脉冲性的，其过程与电晕放电相同。放电结果产生大量的正、负离子，形成空间电荷，建立反电场，使气隙中的总电场强度下降，放电熄灭。这样的放电持续时间很短，为 10^{-8}～10^{-9}s。在直流电压作用时，放电熄灭后直到空间电荷通过表面泄漏，使反电场削弱到一定程度，才能开始第二次放电。因此在直流电压作用下，放电次数甚少。在交流电压作用时，情况就有所不同。由于电压的大小与方向是变动的，放电将反复出现，以上表明局部放电是脉冲性的。

工程介质，从材料本身来说，其本征击穿电场强度一般较高，但由于介质的不均匀性和各种影响，实际击穿强度往往并不很高，有时甚至要降低一、二个数量级，其中重要原因之一就是局部放电。

局部放电将导致介质的击穿和老化，因为局部放电除电的过程以外，还伴随着热、辐射、化学和应力作用等过程。这些过程的综合作用，就使介质击穿或老化变质，破坏过程是相当复杂的。这里介绍一个便于说明、比较简单的破坏机理。大部分陶瓷材料中存在着相当大的气孔，其直径可达几个微米。这些气孔在电场作用下，特别是在高频电场作用下，将发生强烈的游离，而且气孔的直径愈大，游离电压愈低。在高频电压下，由于气孔中的强烈游离，产生大量的热量，使得气孔附近局部区域过热，在材料中产生相当高的内应力。当热应力超过一定限度时，材料因丧失机械强度发生破坏，以致失去抗电能力，造成“击穿”。这种形式的击穿，往往发生在质地疏松、介电常数十分高的材料中。

还有一些其他的击穿机制，如常发生在高分子介质的树枝化击穿和电—机械击穿。

6.4　铁电性

某些电介质不仅具有电介质的共性，还可能有三种特殊性质：压电性、热释电性和铁电性。具有这些特殊性质的电介质，不仅在电子工程中作为传感器、驱

动器材料，还可以在光学、声学、红外探测等领域中发挥独特的作用。这些电介质之所以具有这些特殊性质，完全是由自身的结构决定的。

前面介绍的电介质的极化强度与外加电场的关系，大都是随着电场增加而线性变化，但是一些具有特殊晶体结构（如 $BaTiO_3$）的电介质极化强度随外加电场呈现非线性变化，因此，有人称前者的电介质为线性电介质，而把后者称为非线性电介质。铁电体是一种典型的非线性电介质。

6.4.1 铁电体、电畴

1. 电滞回线和铁电体

1920 年法国人 Valasek 发现罗息盐（酒石酸钾钠—$NaKC_4H_4O_6 \cdot 4H_2O$）具有特异的介电性，其极化强度随外加电场的变化为图 6-17 所示的形状，称为电滞回线，把具有这种性质的晶体称为铁电体。由图 6-17 可见，构成电滞回线的几个重要参量：饱和极化强度 P_s、剩余极化强度 P_r 和矫顽电场 E_c。从电滞回线可以清楚看到铁电体具有自发极化的特性，而且这种自发极化的电偶极矩在外电场作用下可以改变其取向，甚至反转。在同一外电场作用下，极化强度可以有双值，表现为电场 E 的双值函数，这正是铁电体的重要物理特性。铁电体有电滞回线的原因就是存在电畴。

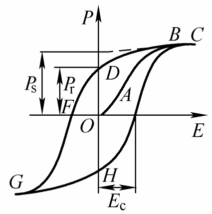


图 6-17 电滞回线

当把罗息盐加热到 24℃以上，则电滞回线便消失，此温度称为居里温度 T_c 或称居里点 T_c。因此，铁电性存在是有一定条件的，包括外界的压力变化。

2. 电畴转向

铁电体整体上呈现自发极化，其结果是晶体正、负端分别有一层正、负束缚电荷。束缚电荷产生的电场—退极化场与极化方向反向，使静电能升高。在受机械约束时，伴随着自发极化的应变还将使应变能增加，所以整体均匀极化的状态不稳定，晶体趋向于分成多个小区域。每个区域内部电偶极子沿同一方向，但不同小区域的电偶极子不同，这每个小区域称为电畴。电畴之间边界地区称之为畴壁。图 6-18 为 $BaTiO_3$ 晶体室温电畴结构示意图。小方格表示晶胞，箭头表示电矩方向。图中 AA 分界线两侧的电矩取反平行方向，称为 180°畴壁，BB 分界线为 90°畴壁。决定畴壁厚度的因素是各种能量平衡的结果，180°畴壁较薄，为（5～20）$\times 10^{-10}$m，而 90°畴壁较厚，为（50～100）$\times 10^{-10}$m。电畴结构与晶体结构有关，一些晶系还有其他角度的畴壁。

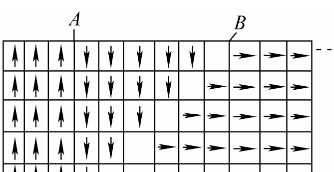


图 6-18 $BaTiO_3$ 晶体电畴结构示意图

铁电畴在外电场作用下，总是趋向与外电场方向一致，称之为畴“转向”。电畴运动是通过新畴出现、发展和畴壁移动来实现的。180°畴转向是通过许多尖劈形新畴出现而发展，90°畴主要是通过畴壁侧向移动来实现。180°转向比较完全，而且由于转向时引起较大内应力，所以这种转向不稳定，当外加电场撤去后，小部分电畴偏离极化方向，恢复原位，大部分电畴则停留在新转向的极化方向上，谓之剩余极化。

电滞回线是铁电体的铁电畴在外电场作用下运动的宏观描述。下面以单晶铁电体为例对前面介绍的电滞回线几个特征参量予以说明。设一单晶体的极化强度方向只有沿某轴的正向或负向两种可能。在没有外电场时，晶体总电矩为零（能量最低）。当加上外电场后，沿电场方向的电畴扩展、变大，而与电场方向反向的电畴变小。这样极化强度随外电场增加而增加，如图 6-17 中的 OA 段。电场强度继续增大，最后晶体电畴都趋于电场方向，类似形成一个单畴，极化强度达到饱和，相应于图中的 C 处。如再增加电场，则极化强度 P 与电场 E 呈线性增加（形如单个弹性电偶极子），沿其线性外推至 $E = 0$ 处，相应的 P_s 值称为饱和极化强度，也就是自发极化强度。若电场强度自 C 处下降，晶体极化强度亦随之减小，在 $E = 0$ 时，仍存在极化强度，就是剩余极化强度 P_r，当反向电场强度为$-E_c$时（图中 F 点处），剩余极化强度 P_r 全部消失；反向电场继续增大，极化强度才开始反向。直到反向极化达到饱和，到达图中 G 处。图中 E_c 称为矫顽电场强度。

由于极化的非线性，铁电体的介电常数不是恒定值，一般以 OA 在原点的斜率来代表介电常数。所以在测定介电常数时，外电场应很小。

6.4.2　钛酸钡自发极化的微观机理及电畴的形成

1. 钛酸钡自发极化的微观机理

关于铁电现象，定量的微观理论还不成熟，下面以钛酸钡为例，介绍其自发极化的微观模型。

由于 $BaTiO_3$，从非铁电相到铁电相的过渡总是伴随着晶格结构的改变，晶体从立方晶系转变为四方晶系，晶体的对称性降低。因此人们提出了一种离子位移理论，认为自发极化主要是由晶体中某些离子偏离了平衡位置造成。由于离子偏离了平衡位置，使得单位晶胞中出现了电矩。电矩之间的相互作用使偏离平衡位置的离子在新的位置上稳定下来，与此同时晶体结构发生了畸变。

钛酸钡的结构为钙钛矿型晶体，Ti^{4+}离子位于 O^{2-}离子八面体中心。在居里温度（120℃）以上，钛酸钡属于等轴晶系，晶胞边长 a =4.01Å（1Å = a /mm），因为 O^{2-}离子半径为 1.32Å，所以两个 O^{2-}离子间的空隙为 4.01−2×1.32 = 1.37Å，而 Ti^{4+}的离子直径为 1.28Å，小于 1.37Å，Ti^{4+}离子在 O^{2-}离子八面体内有位移的余地。在较高温度时（大于 120℃），因为离子热振动能比较大，Ti^{4+}不可能在偏离中心

的某一个位置固定下来，因此它接近周围 6 个 O^{2-}离子的几率是相等的，所以晶体结构仍保持较高的对称性（等轴晶系），晶胞内不会产生电矩，即自发极化为 0；当温度降低时（小于 120℃），Ti^{4+}离子的平均热振动能降低了。那些因热涨落所形成的热振动能量特别低的 Ti^{4+}离子不足以克服 Ti^{4+}和 O^{2-}离子间的电场作用，就有可能向某一个 O^{2-}离子靠近（图 6-19），在此新的平衡位置上固定下来，发生自发位移，并使这个 O^{2-}离子出现强烈的电子位移极化。结果使晶体顺着这个方向延长，晶胞发生轻微的畸变。在 Ti^{4+}离子位移的方向（c 轴）晶轴略有伸长，在其他方向（a、b 轴）缩短，晶体从立方结构转变为四方结构。室温下，c =4.03Å，$a = b = 3.99$Å。结果晶胞中出现了电矩，即发生了自发极化。图 6-20 为离子位移的图形，其数据以中央的四个 O^{2-}离子作为参考点得出。从这些数据可对离子位移引起的极化强度进行估计。一般，自发极化包括两部分：一部分是直接由于离子位移；另一部分是由于电子云的形变。应用洛伦兹表达式可以估计出离子位移极化大约占总极化的 39%。

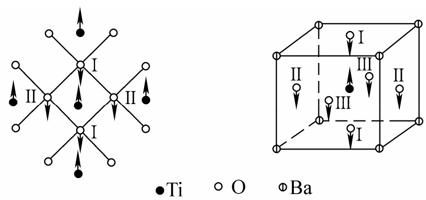


图 6-19　正方结构 $BaTiO_3$ 中钛、氧离子位移情况

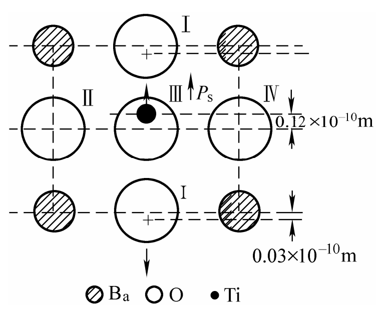


图 6-20　以中央四个 O^{2-}离子为参考的各离子位移情况

以上理论基本是 J. C. Slater 的钛、氧离子强耦合理论。最近关于铁电性的理解已发展成为“软模理论”，它对位移型铁电体是非常成功的。其主要概念是：在一定情况下，晶体结构的铁电相变是布里渊区中心的横向光学晶格振动的“柔软化”结果。所谓柔软化，即声子横光学振动频率接近于零。按量子力学理论，此时振幅无限大，因而晶格结构发生转变，出现铁电相变。

2. 电畴形成

电畴的形成及其运动的微观机理是复杂的。对于 $BaTiO_3$，如果其自发极化的产生由 Ti^{4+}、O^{2-}离子间的强耦合作用引起，则电畴的形成可加以定性解释。如图 6-19，设中间部位的 Ti^{4+}离子因热运动的涨落在某一瞬间向 O^{2-}离子有微小位移，则又使 O^{2-}离子向 Ti^{4+}离子靠拢，接着由于比较大的内电场力的传递，使自发极化首先沿 $Ti^{4+}-O^{2-}$（Ⅰ）离子线展开。同时，由于电场力以及弹性力的传递，周围的 O^{2-}（Ⅱ）离子也被向下挤。如此，自发极化向横向发展。横向发展是间接的，比较弱，因此以上形成的畴核及其发展如针状。最后的电畴图案总是电场力与弹性力平衡的结果，整个体系保持能量最低。铁电畴在外电场作用下，“转向”而趋向与外电场方向一致。

6.4.3 铁电体的临界性质

1. 环境因素对铁电体性质的影响

外部条件（电场、应力、温度、压力）的变化，可以引起铁电体极化强度 P_s 的变化。例如，图 6-21 为铁电体罗息盐的平衡状态图。由图可见，在常压下，它有两个居里点，一是 24℃称上居里点，另一个是−18℃（或−20℃）称为下居里点，在两者中间的温度范围内属于铁电体。铁电体相变按自由能变化来分，可分为两类，即一级相变和二级相变。一级相变时，比热容会发生突变，伴随着有潜热产生，自发极化在居里点处突然下降为零。钛酸钡、PZT 等铁电体即属于这一类。二级相变只呈现比热容大的改变，无潜热发生，自发极化逐渐变为零，硫酸三甘肽（TGS）、$LiTaO_3$ 铁电体属于这一类。图 6-22a、b 分别为它们的自发极化强度与温度的关系曲线。

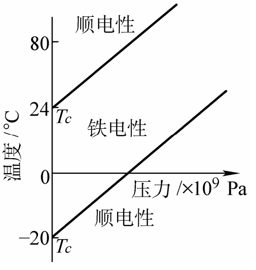


图 6-21　罗息盐相图

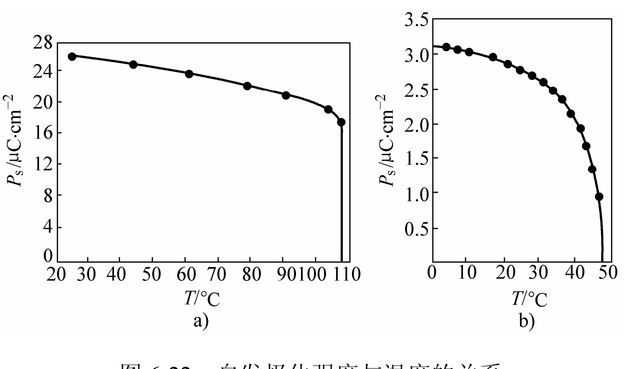


图 6-22　自发极化强度与温度的关系

a）钛酸钡　b）硫酸三甘肽

钛酸钡的自发极化强度 P_s 与温度的关系可由实验得出。120℃以上晶体为立方晶系，无自发极化；120～5℃为四方晶系，自发极化沿 c 轴[001]方向；5～−80℃为斜方晶系，自发极化沿[011]方向；−80℃以下为菱形结构，自发极化沿[111]方向。铁电体中自发极化的突变引起介电常数显著变化，尤其在居里点处。单畴 $BaTiO_3$ 晶体的介电常数与温度的关系如图 6-23 所示。

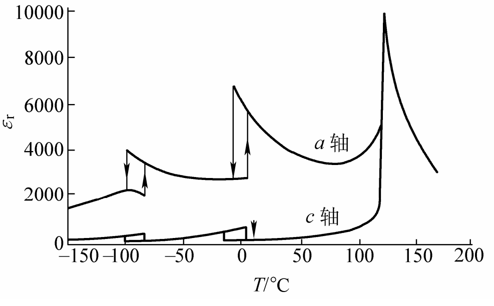


图 6-23 单畴 $BaTiO_3$ 晶体的介电常数与温度的关系（弱场下）

2. 成分、晶粒大小、尺寸因素的影响

铁电体居里温度是由材料成分决定的，不同元素在同一铁电体中对 T_c 的影响是不同的，Pb^{2+}、Sr^{2+}、Ca^{2+}、Cd^{2+}皆可取代 $BaTiO_3$ 中的 Ba^{2+}形成钙钛矿型固溶体，其 T_c 是不同的。同样 Sr^{4+}、Hf^{4+}、Zr^{4+}、Ce^{4+}、Th^{4+}可置换 Ti^{4+}形成铁电体，其 T_c 也不同。

晶粒大小也会影响铁电体的行为。一般情况下晶粒愈大，其压电性（如 d_{33}）值愈高。图 6-24 表示了极细小的晶粒的 $BaTiO_3$ 在 120℃的相变，与 6-23 图比较可见，单晶体相变十分尖锐，而在图 6-24 中的 1～2μm 晶粒的铁电体，相变随温度的变化是逐渐加强的；在极细晶粒情况下（0.2μm），晶粒结构尺寸和钛离子平衡位置间没有或很少有固定的取向关系。因此，在极细晶粒中畴取向是随意的，这种随意性趋于增宽铁电相变的温度范围。

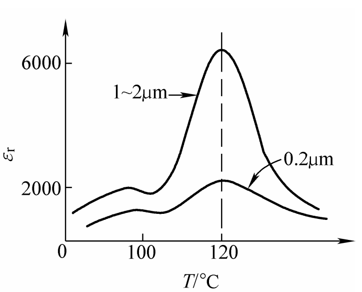


图 6-24 极细晶粒 $BaTiO_3$ 的铁电行为

3. 反铁电—铁电相变

反铁电晶体含有反平行排列的偶极子，$PbZrO_3$的反铁电相结构如图6-25所示。图中箭头方向代表了 Pb^{2+}离子相对于 O^{2-}离子晶格的位移方向，结果在一个正交的晶胞中（实线画出了一个正交晶胞的范围），形成了两个方向相反而偶极矩相等的偶极子亚结构，也就是 $P_1 = P_2$。那么，晶体净极化强度为零。在每个亚晶格结构中，当温度高于 T_c时，极化强度 $P \rightarrow 0$。由于其特殊的偶极子排列，其电滞回线也较特殊，如图 6-26 所示。

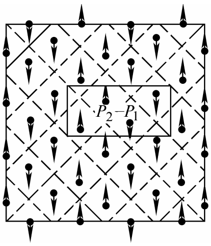


图 6-25　$PbZrO_3$的反铁电相结构

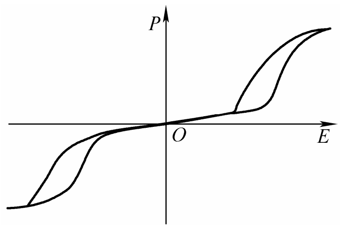


图 6-26　$PbZrO_3$的双电滞回线

反铁电相的偶极子结构很接近铁电相的结构，能量上的差别很小，仅是每摩尔十几焦耳。因此，只要在成分上稍有改变，或者加上强的外电场或者是压力，则反铁电相就转变为铁电相结构。具体实例就是 $PbZrO_3$中的 Zr 为 7%的 Ti 所取代，形成 $Pb(Zr, Ti)O_3$系统，相结构就从反铁电相变成铁电相。

6.5　压电性和热释电性

6.5.1　压电性

压电性就是晶体材料按所施加的机械应力成比例地产生电荷的特性。压电性是 1880 年 P. Curie 和 J. Curie 兄弟发现的，他们对 α—石英单晶体在一些特定方向上加力，则在力的垂直方向的平面上出现正、负束缚电荷，后来称这种现象为压电效应。同时，他们还证实了这类晶体具有可逆的性质。具有压电效应的物体称为压电体。目前已知压电体超千种，它们可以是单晶体、多晶体（如压电陶瓷，各向同性的多晶体需加强电场极化）、聚合物、生物体（如骨骼）。在发明了电荷放大器之后，压电效应获得了广泛应用。

1. 正压电效应、逆压电效应和压电常数

当晶体受到机械力作用时，一定方向的表面产生束缚电荷，其电荷密度大小与所加应力的大小成线性关系。这种由机械能转换成电能的过程，称为正压电效应。正压电效应很早已经用于测力的传感器中。

逆压电效应就是当晶体在外电场激励下，晶体的某些方向上产生形变，而且应变大小与所加电场在一定范围内有线性关系，这种由电能转变为机械能的过程称为逆压电效应。

在正压电效应中，电荷与应力是成比例的，用介质电位移 D（单位面积的电荷）和应力 T 表达如下

$$D = dT \tag{6-84}$$

式中，D 的单位为 $C\cdot m^{-2}$；T 的单位为 $N\cdot m^{-2}$；d 称为压电常数，单位为 $C\cdot N^{-1}$。对于逆压电效应，其应变 S 与电场强度 E（单位 $V\cdot m^{-1}$）的关系为

$$S = dE \tag{6-85}$$

对于正和逆压电效应，比例常数 d 在数值上是相同的，即 $d = D/T = S/E$。实际在以上表示式中，D、E 为矢量，T、S 为张量（二阶对称）。完整地表示压电晶体的压电效应中其力学量（T，S）和电学量（D，E）关系的方程式叫压电方程。采用热力学理论分析，可以导出压电效应相关力学量和电学量之间的定量关系。通常人们以实验方法给出应力与电位移的关系，以便理解正压电效应在晶体上的具体体现。

2. 压电性产生原因

晶体压电效应的本质是因为机械作用（应力与应变）引起了晶体介质的极化，从而导致介质两端表面内出现符号相反的束缚电荷。其机理可用图 6-27 加以解释。图 a 表示压电晶体中质点在某方向上的投影。此时晶体不受外力作用，正电荷重心与负电荷重心重合，整个晶体总电矩为 0，因而晶体表面不荷电。但是当沿某一方向对晶体施加机械力时，晶体由于形变导致正、负电荷重心不重合，即电矩发生变化，从而引起晶体表面荷电；图 b 为晶体在压缩时荷电的情况；图 c 是拉伸时的荷电情况。在后两种情况下，晶体表面电荷符号相反。如果将一块压电晶体置于外电场中，由于电场作用，晶体内部正、负电荷重心产生位移。这一位移又导致晶体发生形变，这个效应即为逆压电效应。

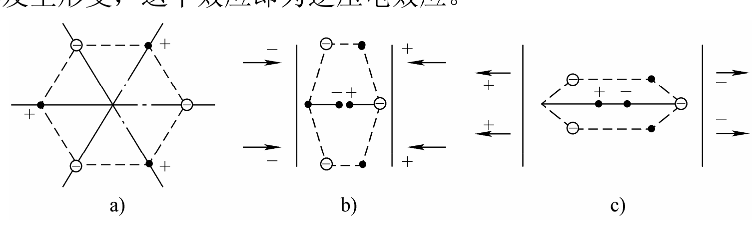


图 6-27　晶体压电效应的示意图

现具体到 α—石英单晶体看其产生压电性的原因。α—石英晶体属于离子晶体三方晶系、无中心对称的 32 点群。石英晶体的化学组成是二氧化硅，3 个硅离子和 6 个氧离子配置在晶胞的格点上。在应力作用下，其两端能产生最强束缚电荷

的方向称为电轴，α—石英的电轴就是 x 轴（图 6-28），z 轴为光轴（光沿此轴进入不产生双折射）。当石英在没有受力的正常态，从 z 轴方向看，则晶胞原子排列如图 6-28a 所示。图中大圆为硅离子，小圆为氧离子。图中硅离子按左螺旋线方向排列，$3^{\#}$硅离子比 $5^{\#}$硅离子较深（向纸内），而 $1^{\#}$硅离子比 $3^{\#}$硅离子更深。每个氧离子带 2 个负电荷，每个硅离子带 4 个正电荷，但每个硅离子的上、下两边有 2 个氧离子，所以整个晶格正、负电荷平衡，不显电性。为了理解正压电效应产生的原因，现把图 6-28a 绘成投影图，把硅离子上、下的氧离子以一个氧离子符号代替并也编成号，如图 6-28b 所示。利用该图定性解释 α—石英晶体产生正压电效应的原因。①如果晶片受到沿 x 方向的压缩力作用，如图 6-28c 所示，这时 $1^{\#}$硅离子挤入 $2^{\#}$和 $6^{\#}$氧离子之间，而 $4^{\#}$氧离子挤入 $3^{\#}$和 $5^{\#}$硅离子之间，结果在表面 A 出现负电荷，而在表面 B 呈现正电荷，这就是纵向压电效应；②当晶片受到沿 y 方向的压缩力作用时，如图 6-28d 所示，这时 $3^{\#}$硅离子和 $2^{\#}$氧离子，以及 $5^{\#}$硅离子和 $6^{\#}$氧离子都向内移动同样数值，故在作用面 C 和 D 上不出现电荷，而在表面 A 和 B 上呈现电荷，但符号与图 6-28c 中的正好相反，因为 $1^{\#}$硅离子和 $4^{\#}$氧离子向外移动。这称之为横向压电效应；③当沿 z 方向压缩或拉伸时，带电粒子总是保持初始状态的正、负电荷重心重合。故表面不出现束缚电荷。

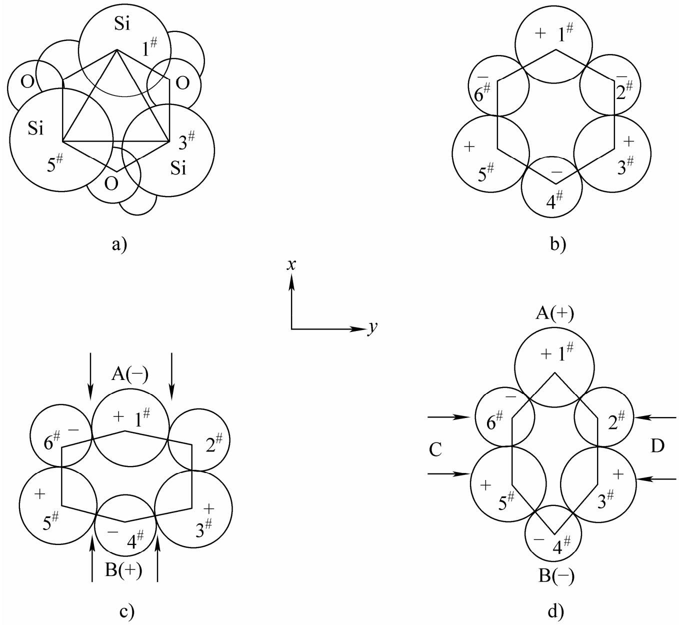


图 6-28　α—石英产生正压电效应的示意图

一般情况下正压电效应的表现是晶体受力后在特定平面上产生束缚电荷。但其直接作用力使晶体产生应变，即改变了原子相对位置。产生束缚电荷的现象，

表明出现了净电偶极矩。如果晶体结构具有对称中心，那么只要作用力没有破坏其对称中心结构，则正、负电荷的对称排列，也不会改变，即使应力作用产生应变，也不会产生净电偶极矩。因为具有对称中心的晶体总电矩为零。如果取一无对称中心的晶体结构，此时正、负电荷重心重合，加上外力后，正负电荷重心不再重合，结果产生净电偶极矩。因此，从晶体结构上分析，只要结构没有对称中心，则就有可能产生压电效应，然而并不是没有对称中心的晶体一定是有压电性，因为压电体首先必须是电介质（或至少具有半导体性质），同时其结构必须有带正、负电荷的质点—离子或离子团存在。也就是说压电体必须是离子晶体或者由离子团组成的分子晶体。

3. 压电振子及其参数

压电振子是最基本的压电元件，它是被覆激励电极的压电体。样品的几何形状不同，可以形成各种不同的振动模式。表征压电效应的主要参数，除以前讨论的介电常数、弹性模量和压电常数等压电材料的常数外，还有表征压电元件的参数，这里重点讨论谐振频率、机械品质因数、频率常数和机电耦合系数。

（1）谐振频率与反谐振频率

若压电振子是具有固有振动频率 f_r 的弹性体，当施加于压电振子上的激励信号频率等于 f_r 时，压电振子由于逆压电效应产生机械谐振，这种机械谐振又借助于正压电效应而输出电信号。压电振子谐振时，输出电流达最大值，此时的频率为最小阻抗频率 f_m。当信号频率继续增大到 f_n，输出电流达最小值，f_n 叫最大阻抗频率，如图 6-29 所示。

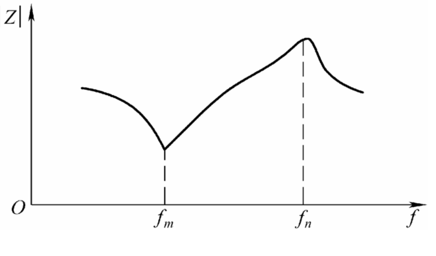


图 6-29 压电振子的阻抗特性曲线示意图

根据谐振理论，压电振子在最小阻抗频率 f_m 附近，存在一个使信号电压与电流同位相的频率，这个频率就是压电振子的谐振频率 f_r，同样在 f_n 附近存在另一个使信号电压与电流同位相的频率，这个频率叫压电振子的反谐振频率 f_a。只有压电振子在机械损耗为零的条件下，$f_m = f_r$，$f_n = f_a$。

（2）机械品质因数

压电振子谐振时，仍存在内耗，造成机械损耗，使材料发热，降低性能。反映这种损耗程度的参数称为机械品质因数 Q_m，其定义式为

$$Q_m = 2\pi \frac{W_m}{\Delta W_m} \tag{6-86}$$

式中，W_m 为振动一周单位体积存储的机械能；ΔW_m 为振动一周内单位体积内消耗的机械能。不同压电材料的机械品质因数 Q_m 的大小不同，而且还与振动模式有关。

（3）频率常数

压电元件的谐振频率与沿振动方向的长度的乘积为一常数，称为频率常数 N（单位 kHz·m）。例如陶瓷薄长片沿长度方向（长度为 l）伸缩振动的频率常数 $N_l = f_r l$，而 $f_r = [1/(2l)](Y/\rho)^{1/2}$，这里 Y 为弹性模量，ρ 为材料密度。所以有

$$N_l = \frac{1}{2}\sqrt{\frac{Y}{\rho}} \tag{6-87}$$

由此可见，频率常数只与材料的性质有关。若知道材料的频率常数即可根据所要求的频率来设计元件的外形尺寸。

（4）机电耦合系数

机电耦合系数 k 是综合反映压电材料性能的参数。它表示压电材料的机械能与电能的耦合效应，定义为

$$k^2 = \frac{\text{由机械能转化的电能}}{\text{输入的总机械能}} \tag{6-88}$$

或

$$k^2 = \frac{\text{由电能转化的机械能}}{\text{输入的总电能}} \tag{6-89}$$

由于压电元件的机械能与它的形状和振动方式有关，因此不同形状和不同振动方式所对应的机电耦合系数也不相同。

4. 压电陶瓷的预极化

所谓极化，就是在压电陶瓷上加一个强直流电场，使陶瓷中的电畴沿电场方向取向排列。只有经过极化工序处理的陶瓷，才能显示压电效应。这里从压电性能出发，讨论极化的条件。

（1）极化电场

极化电场是极化诸条件中的主要因素。极化电场越高，促使电畴取向排列的作用越大，极化就越充分。一般以平面机电耦合系数 k_p 达到最大值的电场为极化电场，但应注意，不同的机电耦合系数达到最大值的极化电场不一样。极化电场必须大于样品的矫顽场，通常为矫顽场的二、三倍。矫顽场与样品的成分、结构及温度有关。

（2）极化温度

在极化电场和时间一定的条件下，极化温度高，电畴取向排列较易，极化效果好。这可从两方面理解：①结晶各向异性随温度升高而降低，自发极化重新取

向克服的应力阻抗较小。同时由于热运动，电畴运动能力加强；②温度越高，电阻率越小，由杂质引起的空间电荷效应所产生的电场屏蔽作用小，故外加电场的极化效果好，但是温度过高，击穿强度降低，常用压电陶瓷材料的极化温度通常取 320～420K。

（3）极化时间

极化时间长，电畴取向排列的程度高，极化效果较好。极化初期主要是 180°电畴的反转，以后的变化是 90°电畴的转向。90°电畴转向由于内应力的阻碍而较难进行，因而适当延长极化时间，可提高极化程度。一般极化时间从几分钟到几十分钟。

总之，极化电场，极化温度，极化时间三者必须统一考虑，因为它们之间相互有影响，应通过实验选取最佳条件。

6.5.2 热释电性

一些晶体除了由于机械应力作用引起压电效应外，还可以由于温度作用而使其电极化强度变化，这就是热释电性，亦称热电性。

1. 热释电现象

将一块电气石（化学组成为$(Na, Ca)(Mg, Fe)_3B_3A1_6Si_6(O, OH, F)_{31}$）在均匀加热它的同时，让一束硫磺粉和铅丹粉经过筛孔喷向这个晶体，结果会发现，晶体一端出现黄色，另一端变为红色。这就是坤特法显示的天然矿物晶体电气石的热释电性实验。实验表明，如果电气石不是在加热过程中，喷粉实验不会出两种颜色的。现在已经认识到，电气石是三方晶系 3m 点群。结构上只有惟一的三次（旋）转轴，具有自发极化。没有加热时，它们的自发极化电偶极矩完全被吸附的空气中的电荷屏蔽掉了。但在加热时，由于温度变化，使自发极化改变，则屏蔽电荷失去平衡。因此，晶体一端的正电荷吸引硫磺粉显黄色，另一端吸引铅丹粉显红色。这种由于温度变化而使极化改变的现象称热释电效应，其性质称为热释电性。

2. 热释电效应产生的条件

热释电效应研究表明，具有热释电效应的晶体一定是具有自发极化（固有极化）的晶体，在结构上应具有极性轴，简称极轴。所谓极轴，就是晶体惟一的轴，在该轴二端往往具有不同性质，且采用对称操作不能与其他晶向重合的方向，谓之极轴。因此，具有对称中心的晶体是不可能有热释电性的，这一点与压电体的结构要求是一样的。但具有压电性的晶体不一定有热释电性。原因可以从二者产生的条件来分析：当压电效应发生时，机械应力引起正、负电荷的重心产生相对位移，而且一般说不同方向上位移大小是不相等的，因而出现净电偶极矩。而当温度变化时，晶体受热的膨胀却在各方向同时发生，并且在对称方向上必定有相等的膨胀系数，也就是说在这些方向上所引起的正、负电荷重心的相对位移也是

相等的，也就是正、负电荷重心重合的现状并没有因为温度变化而改变，所以没有热释电现象。

表征材料热释电性的主要参量是热释电常量 p，其定义为

$$p = \frac{\partial P_s}{\partial T} \tag{6-90}$$

式中，P_s 为自发极化强度；T 为热力学温度。

3. 铁电性、压电性、热释电性之间的关系

至此，已经介绍了一般电介质、具有压电性的电介质（压电体）、具体热释电性的电介质（热电体）、具有铁电性的电介质（铁电体），它们存在的宏观条件如表 6-3 所示。它们的关系如图 6-30 所示。从此图可见，铁电体一定是压电体和热释电体。在居里温度以上，有些铁电体已无铁电性，但其顺电体仍无对称中心，故仍有压电性，如磷酸二氢钾。有些顺电相如钛酸钡是有对称中心的，故在居里温度以上既无铁电性也无压电性，总之，与它们的晶体结构密切相关。

表 6-3　一般电介质、压电体、热电体、铁电体存在的宏观条件

电介质	压电体	热电体	铁电体
电场极化	电场极化	电场极化	电场极化
	无对称中心	无对称中心	无对称中心
		自发极化	自发极化
		极轴	极轴
			电滞回线

注：有学者认为，铁电体不一定有完整的电滞回线，只要在外电场作用下自发偶极矩可改变方向即可。

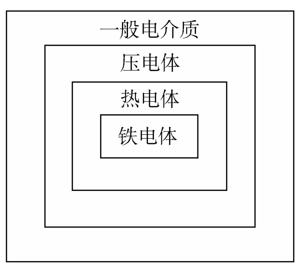


图 6-30　一般电介质、压电体、热电体、铁电体之间的关系

第 7 章　材料的磁性能

磁性与物质的微观结构密切相关，它不仅取决于物质的原子结构，还取决于原子间的相互作用—键合情况和晶体结构。因此，研究磁性是研究物质内部结构的重要方法之一，也为发现新型磁性材料奠定基础。

本章将着重介绍原子磁矩，物质磁性的分类。进一步介绍铁磁性的分子场理论，亚铁磁性的超交换理论。也介绍了铁磁性物质内部的能量和磁畴的形成。

7.1　磁学基本量及物质磁性分类

7.1.1　磁学基本量

1. 磁矩

一个圆电流的磁矩可定义为

$$m = IS \qquad (7\text{-}1)$$

式中，I 是圆电流的强度（A）；S 是圆电流回路所包围的面积（m^2）；m 为圆电流的磁矩，单位为 $A{\cdot}m^2$，方向可用右手定则来确定。

磁性材料单位体积的磁矩称为磁化强度，用 M 表示，单位为 $A{\cdot}m^{-1}$。

2. 磁场强度

一无限长螺旋线圈管，每米有 N 匝，通过电流为 I，管中心轴线上的磁场强度大小的定义为

$$H = NI \qquad (7\text{-}2)$$

磁场强度的单位为 $A{\cdot}m$。

3. 磁感应强度

在磁场中，单位速率的单位电荷所受的最大磁力定义为磁感应强度，用 B 表示，单位为 T。

实践和理论都已证明，任何物质在外磁场的作用下，除了外磁场外，由于物质内部的微观电流，还要产生一个附加的磁场。外磁场和附加磁场的和称为材料的磁感应强度，用 B 表示。在真空中，磁感应强度 B 与外磁场 H 成正比，即

$$B = \mu_0 H \qquad (7\text{-}3)$$

式中，μ_0 为真空磁导率，$\mu_0= 4\pi\times10^{-7}H{\cdot}m^{-1}$。将材料放入磁场为 H 的自由空间中，材料的磁感应强度 B 为

$$B=\mu_0(H+M) \tag{7-4}$$

4. 磁位能（静磁能）

位于均匀磁场中的磁铁，当磁矩方向与磁场方向不一致时，磁铁将受到一个力矩作用，力矩力图使磁铁转动到与磁场一致的方向。单位体积磁铁所受力矩的大小为

$$L=MH\sin\theta \tag{7-5}$$

式中，M 是磁铁的磁化强度；H 为外加磁场的强度；θ 为磁铁内部自己的磁场方向（由 S 极指向 N 极）与外加磁场方向的夹角。

在磁力矩的作用下，力图使 $\theta=0$。如果外力推动磁铁，使 θ 角增加 $\mathrm{d}\theta$，则外力要反抗磁力矩做功，从而使磁铁单位体积能量增加 $\mathrm{d}E_H=L\mathrm{d}\theta$。当磁铁在磁场中转动 θ 角时，其单位体积能量增加

$$E'_H=\int L\mathrm{d}\theta=\int MH\sin\theta\mathrm{d}\theta \tag{7-6}$$

假定 M 和 H 是均匀的，则 $E'_H=-MH\cos\theta+C$。设 $\theta=90°$时，$E'_H=0$，$C=0$，则磁铁在磁场中的磁位能为

$$E_H=-MH\cos\theta \tag{7-7}$$

式中，E_H 也称为静磁能，单位为 $\mathrm{J\cdot m^{-3}}$。

5. 磁化率和磁导率

体积磁化率 χ（磁化率）和绝对磁导率 μ 定义为

$$\left.\begin{aligned}\chi&=M/H\\ \mu&=B/H\end{aligned}\right\} \tag{7-8}$$

绝对磁导率 μ 除以真空磁导率 μ_0，则得相对磁导率 $\mu_r=\mu/\mu_0$，μ 的单位为 $\mathrm{H\cdot m^{-1}}$。χ 和 μ_r 是无量纲的，它们的关系为

$$\mu_r=1+\chi \tag{7-9}$$

磁性参量 χ、μ_r、μ 实质上是描写同一客观现象，已知其中一个，就可以确定其他两个。

7.1.2　物质的磁性分类

根据物质磁化率，可以把物质的磁性大致分为五类。按各类磁体磁化强度 M 与磁场强度 H 的关系，可做出其磁化曲线。图 7-1 为它们的磁化曲线示意图。

1. 抗磁体

磁化率 χ 为甚小的负数，大约在 10^{-6} 数量级。它们在磁场中受微弱斥力。金属中约有一半简单金属是抗磁体。根据 χ 与温度的关系，抗磁体又可分为：①“经

典”抗磁体，它的 χ 不随温度变化，如铜、银、金、汞、锌等；②反常抗磁体，它的 χ 随温度变化，且其大小是前者的 10～100 倍。如铋、镓、锑、锡、铟、铜锆合金中的 γ 相等。

2. 顺磁体

磁化率 χ 为正值，约为 10^{-6}～10^{-3}。它在磁场中受微弱吸力。又根据 χ 与温度的关系可分为：①正常顺磁体，其 χ 随温度变化。符合 $\chi\propto 1/T$ 关系（T 为温度）。金属铂、钯、奥氏体不锈钢、稀土金属等属于此类；②χ 与温度无关的顺磁体，例如锂、钠、钾、铷等金属。

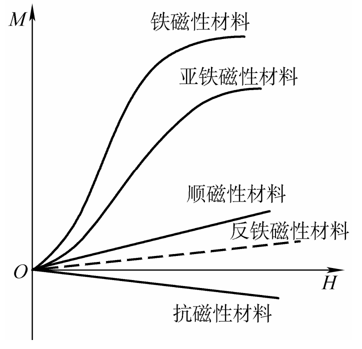


图 7-1　五类磁体的磁化曲线示意图

3. 铁磁体

在较弱的磁场作用下，就能产生很大的磁化强度。χ 是很大的正数，且与外磁场呈非线性关系变化。具体金属有铁、钴、镍等。铁磁体在温度高于某临界温度后变成顺磁体。此临界温度称为居里温度或居里点，常用 T_c 表示。

4. 亚铁磁体

这类磁体有些像铁磁体，但 χ 值没有铁磁体那样大。通常所说的磁铁矿（Fe_3O_4）、铁氧体等属于亚铁磁体。

5. 反铁磁体

这类磁体的 χ 是小的正数，在温度低于某温度时，它的磁化率同磁场的取向有关；高于这个温度，其行为像顺磁体。具体材料有 α-Mn、铬，还有如氧化镍、氧化锰等。

7.2　原子和离子的固有磁矩

原子的磁矩来源于电子和原子核，但原子核的磁矩仅有电子磁矩的 1/1836.5，所以原子磁矩主要来源于电子磁矩。下面以玻尔和量子力学的原子结构理论为基础，介绍原子的磁矩。

7.2.1　孤立原子（离子）本征磁矩

1. 电子轨道磁矩

电子绕原子核运动，犹如一环形电流，此环流也应在其运动中心处产生磁矩，称为电子轨道磁矩。设电子轨道半径为 r，绕轨角速度为 ω。电子的电量为 e，质量为 m，电子的角动量矩 $P_l = mr^2\omega$。根据磁矩等于电流与电流回路所包围的面积

的乘积的原理，可得电子的轨道磁矩 μ_l 为

$$\mu_l = -\frac{e}{2mc}P_l \tag{7-10}$$

式中，c 为光速；负号表明电子轨道磁矩 μ_l 与轨道动量矩 P_l 的方向相反。由玻尔量子轨道理论知道，在同一主量子层 n 上的电子轨道动量矩 $P_l = n_l\hbar$，这里 n_l =1, 2, 3, …, n，称为角量子数；$\hbar = h/(2\pi)$，h 为普朗克常数。代入式（7-10），且去掉代表方向的符号得

$$\mu_l = \frac{e\hbar}{2mc}n_l \tag{7-11}$$

当主量子数 n =1，n_l =1 时，即 1s 电子轨道磁矩为

$$\mu_l = \mu_B = \frac{e\hbar}{2mc} \tag{7-12}$$

μ_B 称为玻尔磁子，是磁矩的最小单位，$\mu_B = 9.2732\times10^{-24}\text{A}\cdot\text{m}^2$。

上述推导是根据玻尔的经典原子结构模型进行的，但是准确的电子运动要用量子力学中的波函数来描述。利用量子力学的结果，与电子轨道运动相对应的电子角动量矩 P_l 为

$$P_l = \sqrt{l(l+1)}\hbar \tag{7-13}$$

式中，l 可取 0，1，2，…，(n–1)，它们分别代表 s，p，d，f，…电子态，称为轨道角量子数。n 为主量子数。将式（7-13）代入式（7-10），且去掉代表方向的符号得电子轨道磁矩 μ_l 为

$$\mu_l = \sqrt{l(l+1)}\mu_B \tag{7-14}$$

根据量子理论，在外电场作用下，轨道磁矩在外磁场方向的投影值 $\mu_{l\cdot H}$ 并不能任意取值，而仅能取值

$$\mu_{l\cdot H} = m_l\mu_B \tag{7-15}$$

式中，m_l 为磁量子数，m_l = 0，±1，±2，±3，…，±l，共（2l+1）个值。说明电子轨道磁矩在磁场中的投影值是量子化的。

在填满了电子的次电子层（s、p、d、f、…）中，各电子的轨道运动分别占据了所有可能的方向，形成一个球形对称体系，因此合成的总轨道角动量等于零，总轨道磁矩也等于零。例如 3d 态电子，n =3，l =2，如果 3d 层填满了 10 个电子，则这 10 个电子轨道磁矩在磁场方向的投影总和为$[0+1+2+(-1)+(-2)]\times2\,\mu_B = 0$。所以计算原子的总轨道磁矩时，只需要考虑未填满的那些次壳层中电子的贡献。

2. 电子自旋磁矩

电子自旋运动是由于量子力学效应，在宏观物体中还找不出一种运动与之相对应。实验和量子力学已证明电子在作轨道运动的同时还绕自身的轴做自旋运动，

自旋运动产生的磁矩 μ_s 为

$$\mu_s = 2\sqrt{s(s+1)}\mu_B \tag{7-16}$$

式中，s 为自旋量子数，它仅能取 1/2。自旋磁矩在外磁场中的投影值 $\mu_{s\cdot H}$ 为

$$\mu_{s\cdot H} = 2m_s\mu_B \tag{7-17}$$

式中，$m_s= \pm 1/2$，称为自旋角动量方向量子数。如果 s、p、d、f 等次电子层填满了电子时，电子总自旋磁矩也为零。所以计算原子的总自旋磁矩时，只需要考虑未填满的那些次壳层中电子的贡献。

3. 原子的总磁矩

原子磁矩应是电子轨道磁矩与自旋磁矩的总和。但电子轨道磁矩和电子自旋磁矩是如何耦合成原子总磁矩呢？有两种可能性。第一是 *JJ* 耦合。当原子序数大于 82 时，各电子轨道磁矩与自旋磁矩作用较强。各个电子的轨道磁矩和自旋磁矩耦合成电子的总磁矩后，再由各个电子的总磁矩耦合成原子的总磁矩，这种耦合称为 *JJ* 耦合。第二是 *LS* 耦合。当原子序数小于 32 时，由于轨道磁矩与自旋磁矩的相互作用较弱，而电子的轨道磁矩间，自旋磁矩间作用较强。因此原子内各电子轨道磁矩先组合成原子总的轨道磁矩，各电子的自旋磁矩先组合成原子总的自旋磁矩，然后两者再耦合成原子的总磁矩，这种耦合称为 *LS* 耦合。原子序数介于 32 和 82 间的原子，两种耦合都存在，随原子序数增加，先由 *LS* 耦合为主，逐渐过渡到以 *JJ* 耦合为主。在 3d 过渡族金属和 4f 稀土金属及合金中，主要是 *LS* 耦合。

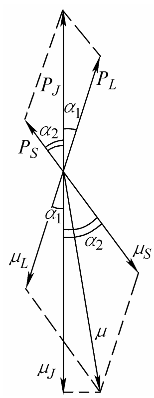


图 7-2 $\boldsymbol{\mu}_L$ 和 $\boldsymbol{\mu}_S$ 矢量图

现在讨论 *LS* 耦合。设原子内各电子轨道磁矩组合成原子总的轨道磁矩为 $\boldsymbol{\mu}_L$，各电子轨道角动量组成原子总的轨道角动量为 $\boldsymbol{P}_L$；各电子的自旋磁矩组合成原子总的自旋磁矩为 $\boldsymbol{\mu}_S$，各电子的自旋角动量组成原子总的自旋角动量为 $\boldsymbol{P}_S$；原子磁矩为 $\boldsymbol{\mu}_J$，原子角动量为 $\boldsymbol{P}_J$（$\boldsymbol{P}_J=\boldsymbol{P}_L+\boldsymbol{P}_S$），如图 7-2 所示。原子磁矩 $\boldsymbol{\mu}_J \neq \boldsymbol{\mu} = \boldsymbol{\mu}_L + \boldsymbol{\mu}_S$，$\boldsymbol{\mu}_J$ 等于 $\boldsymbol{\mu}_L$ 和 $\boldsymbol{\mu}_S$ 在 $-\boldsymbol{P}_J$ 投影之和（图 7-2），即

$$\begin{aligned}\mu_J &= \mu_L\cos\alpha_1 + \mu_S\cos\alpha_2 \\ &= \sqrt{L(L+1)}\mu_B\cos\alpha_1 + 2\sqrt{S(S+1)}\mu_B\cos\alpha_2\end{aligned} \tag{7-18}$$

式中，L 为原子总轨道角量子数；S 为原子总自旋量子数（根据洪德法则计算，见后）；α_1 为 $\boldsymbol{P}_L$ 与 $\boldsymbol{P}_J$ 的夹角；α_2 为 $\boldsymbol{P}_S$ 与 $\boldsymbol{P}_J$ 的夹角。式（7-18）简化后得

$$\mu_J = g_J\sqrt{J(J+1)}\mu_B \tag{7-19}$$

式中，J 为原子总角量子数（根据洪德法则计算，见后）；g_J 为朗德因子，与 J、

L 和 S 的关系为

$$g_J = 1 + \frac{J(J+1) + S(S+1) - L(L+1)}{2J(J+1)} \tag{7-20}$$

当 $L=0$ 时，$J=S$，$g_J=2$，原子总磁矩由自旋磁矩贡献。当 $S=0$ 时，$J=L$，$g_J=1$，原子总磁矩由轨道磁矩贡献。

量子力学理论还证明，原子总磁矩在外加磁场方向的投影 $\mu_{J\cdot H}$ 为

$$\mu_{J\cdot H} = g_J m_J \mu_B \tag{7-21}$$

式中，m_J 是原子的磁量子数，$m_J = J$，$J-1$，$J-2$，…，$-J$，共（$2J+1$）个值。当 m_J 取最大值 J 时，得到原子磁矩在磁场方向的最大分量：$(\mu_{J\cdot H})_{max}=g_J J\mu_B$，它和物质的磁性有密切关系。例如在单位体积中有 n 个磁性原子的体系，它在 0K 时自发磁化强度 $Ms(0) = ng_J J\mu_B$。

4. 洪德（Hund）法则与孤立基态原子或离子磁矩的计算

洪德（Hund）根据原子光谱实验，总结了计算基态原子或离子的总角量子数 J 的法则，称洪德法则。其主要内容为：在 LS 耦合的情况下，对那些次电子层未填满电子的原子或离子，在基态下，其总角量子数 J 与总轨道量子数 L 和总自旋量子数 S 的取值为：①在未填满电子的那些次电子层内，在泡利（Pauli）原理允许的条件下，总自旋量子数 S 取最大值，总轨道量子数 L 也取最大值；②次电子层未填满一半时，原子总角量子数 $J= L-S$；次电子层满一半或满一半以上时，原子的总角量子数 $J= L+S$。

根据洪德法则，可以计算孤立基态原子或离子的磁矩。例如三价镝离子 Dy^{3+} 的电子排布为 $1s^2\ 2s^2\ 2P^6\ 3s^2\ 3p^6\ 3d^{10}\ 4s^2\ 4p^6\ 4d^{10}\ 4f^9\ 5s^2\ 5p^6$。可见除了 4f 次电子层外，其他次电子层已填满了电子，它们的轨道磁矩与自旋磁矩对原子磁矩没有贡献，因此仅计算 4f 次电子层的磁矩就可以了。4f 电子层的 $n=4$，$l=3$，有 9 个电子，已过半满。这 9 个电子的电子态及相应的量子数如表 7-1 所示。根据洪德法则得

表 7-1　Dy^{3+}中 4f 电子态的电子分布及相应的量子数

m_l	电子数	自旋方向	m_s
+3	1+1	↑↓	±1/2
+2	1+1	↑↓	±1/2
+1	1	↑	+1/2
0	1	↑	+1/2
−1	1	↑	+1/2
−2	1	↑	+1/2
−3	1	↑	+1/2

$$\left.\begin{aligned} L &= (\sum m_l)_{\max} = 3+3+2+2+1+0+(-1)+(-2)+(-3) = 5 \\ S &= (\sum m_s)_{\max} = 7\times\frac{1}{2}+2\times\left(-\frac{1}{2}\right)=\frac{5}{2} \\ J &= L+S = 5+\frac{5}{2}=\frac{15}{2} \end{aligned}\right\} \quad (7\text{-}22)$$

将 L、S、J 代入式（7-19）和式（7-20）得 $g_J = 1.333$，$\mu_J = 10.64\mu_B$。由式$(\mu_{J\cdot H})_{\max} = g_J J\mu_B$ 得$(\mu_{J\cdot H})_{\max} = 10.00\mu_B$。

7.2.2　固体中的原子（离子）磁矩

上面叙述了孤立、自由的原子（离子）的磁矩。那么原子组成物质时，根据材料结构不同，物质中的原子磁矩有时会发生变化。

1. 铁氧体中的原子磁矩

在主要由 Fe 等 3d 过渡族离子组成铁氧体晶体时，其特点是对磁性有贡献的 3d 电子基本固定在原子周围，它受到了邻近离子原子核的库仑场以及电子的作用，这一作用的平均效果等效为一个电场，称为晶体场。3d 电子的主量子数 $n = 3$，角量子数 l=2，它有 5 种波函数，$\psi_1(x^2-y^2)$，$\psi_2(z^2-x^2-y^2)$，$\psi_3(xy)$，$\psi_4(yz)$，$\psi_5(zx)$和磁量子数 $m_l = \pm1$，±2，0 对应。虽然 m_l 不同，但是在自由原子（离子）的情况，原子的能量仅和主量子数 n、角量子数 l 有关，所以 5 种轨道处于相同的能级，称能量简并，如图 7-3a 所示。但是在晶体中情况就会不同，例如 Fe_3O_4 晶体中，Fe^{2+} 处于一个八面体中心，负离子在八面体的各个顶点，负离子和它的相对位置如图 7-3b 所示。很明显，$\psi_1(x^2-y^2)$和 $\psi_2(z^2-x^2-y^2)$，轨道受负离子的电场影响最大，负

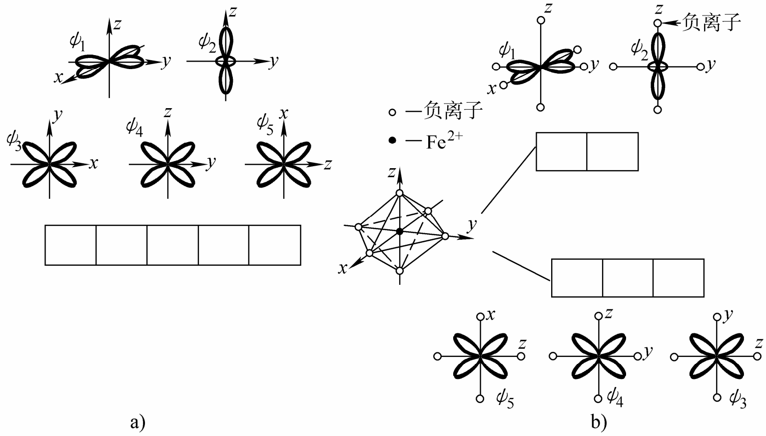


图 7-3　Fe^{2+}在八面体中的能级分裂

a）孤立原子的能级　b）Fe^{2+}在八面体中的能级

离子和电子相互排斥，使轨道的能量增加，结果 $\psi_1(x^2-y^2)$和 $\psi_2(z^2-x^2-y^2)$轨道的能量变得高于其他 3 个轨道的能量，称为能级的简并部分消除，如图 7-3b 所示，导致轨道运动对磁矩的贡献降低。

例如有一个 3d 电子（$3d^1$）的原子，在自由原子时：$L = 2$，$S = 1/2$，$J = 3/2$。该原子处于八面体晶场中时，由相应的计算可以推出 $L = 1$，$S = 1/2$，因此 $J = 1/2$。又因为$(\mu_{J\cdot H})_{\max} = g_J J\mu_B$，所以该原子处在八面体晶场中 $\mu_{J\cdot H}$ 比自由原子的 $\mu_{J\cdot H}$ 小。这里原子受晶体场影响使得 $L=1$，小于它为自由原子时的 $L = 2$，称轨道部分冻结。如果受晶体场影响使得 $L = 0$，则称轨道完全冻结，这时轨道运动对磁矩就没有贡献。

在铁氧体等磁性材料中，计算原子磁矩时一般近似认为轨道完全冻结，只考虑它的自旋磁矩的贡献。这样计算出来的原子磁矩和实验值很吻合。例如 $NiFe_2O_4$ 铁氧体，其中一个 Fe^{3+}离子和另一个 Fe^{3+}离子、Ni^{2+}离子的磁矩是反平行排列，在只考虑自旋时，Fe^{3+}的 $\mu_{J\cdot H} = g_J J\mu_B = 2S\mu_B = 5\mu_B$，$Ni^{2+}$的 $\mu_{J\cdot H} = 2S\mu_B = 2\mu_B$。这样分子磁矩 $m = (2+5-5)\,\mu_B = 2\mu_B$，而实验值为 $2.3\mu_B$，非常吻合。

2. 金属晶体中的原子（离子）磁矩

大块金属的原子磁矩 μ_J 是先测量大块金属的饱和磁化强度 M_s，再用 M_s 除以单位体积的原子数得到的。

（1）镧系金属晶体的原子（离子）磁矩

上面计算了孤立镝离子 Dy^{3+}的原子磁矩，理论计算值 $\mu_{J(理)} = 10.64\mu_B$，而大块金属镝的原子磁矩实验值 $\mu_{J(实)} = 10.63\mu_B$，两者相符很好。整个镧系（孤立）原子磁矩理论值与实验值（当然实验值均为大块金属的）都相符的很好，说明在镧系金属（稀土）中自由原子（离子）磁矩与固态金属中原子磁矩是相同的。

（2）3d 过渡族金属晶体中原子（或离子）磁矩

3d 过渡族金属原子磁矩的理论值与实验值相差甚远。如 Ni 金属 $\mu_{J(理)}= 5.58\mu_B$，$\mu_{J(实)} = 0.606\mu_B$。又如 Co 金属 $\mu_{J(理)} = 6.4\mu_B$，$\mu_{J(实)} = 1.716\mu_B$。3d 过渡族金属和 4f 稀土金属大不相同。在 4f 金属中，对磁性有贡献的是 4f 电子。在 4f 电子之外还有 5s 和 5p 电子层，起到屏蔽作用，使 4f 电子不受晶场的影响，其轨道磁矩均有贡献。但是在 3d 金属晶体中，对原子磁矩有贡献的 3d 电子是最外层电子，它强烈地受到晶格场的影响，其轨道磁矩被晶场所控制，不能随外磁场转动，对原子磁矩没有贡献。这种现象称为轨道磁矩“冻结”。而对原子磁矩有贡献仅是 3d 电子的自旋磁矩。这样，根据式（7-21），Fe、Co 和 Ni 的原子磁矩应分别是 $4\mu_B$、$3\mu_B$、$2\mu_B$。然而实验值仍然比这些数值小，并且不是整数。这一点可用能带理论来说明。根据能带理论，它的最外层 4s 电子是自由电子，可以在晶体中自由移动，组成金属的各个原子的 4s 电子轨道完全重合，4s 能级变成了很宽的能带，4s 电子已经不属于哪个原子了。同样，次外层的 3d 电子虽说不像 4s 电子那样可以自由

地在晶体中移动，但是它也不是完全局域在某个原子周围。在一定程度上，3d电子也可以自由移动，它的能级也变成了能带，并且和4s能带重叠。因此具有同样能量的电子可以进入3d轨道，也可以进入4s轨道，使得金属Fe、Ni及其合金的磁矩和将3d电子完全考虑为某个孤立原子的电子不同。所以要从能带理论的角度来解释这些材料的原子磁矩。如果3d+4s的电子数超过8时，可以用经验公式来计算Fe、Ni金属及其合金的原子磁矩 $\mu_{J\cdot H} = (10.6-n)\,\mu_B$，这里 n 是4s+3d电子数。由此计算出Fe（$n=8$）、Ni（$n=10$）、Co（$n=9$）的 $\mu_{J\cdot H}$ 分别为 $2.6\mu_B$、$0.6\mu_B$、$1.6\mu_B$，和实验值比较接近。

7.3 物质的抗磁性和顺磁性

前面已经讲到，根据物质磁化率，可以把物质的磁性大致分为五类，其中有抗磁体和顺磁体。在这里主要讲抗磁性和顺磁性的物理本质。

7.3.1 抗磁性

抗磁性来源于将物质放入外磁场中时，穿过电子轨道运动回路的磁通会发生变化，为了抵抗该变化，在电子轨道回路要产生一个附加的感应电流（由轨道半径发生变化造成）。这一附加感应电流的磁矩方向和外磁场方向相反，因此抗磁性的磁化率 χ 是负的。基特尔（Kittel）根据经典电动力学的原理得出抗磁磁化率 χ 为

$$\chi = -\frac{nz\mu_0 e^2}{6m} r^2 \tag{7-23}$$

式中，n 为单位体积的原子数；m 为电子的静止质量；e 为电子的电荷；r^2 是电子云半径平方的平均值；z 是原子的电子数；μ_0 是真空磁导率。由量子力学理论也可得到式（7-23）。对于惰性气体，由式（7-23）算得的磁化率理论值与实验值符合得很好。如氦气，摩尔磁化率为 -2.39×10^{-11}，其实验值为 -2.40×10^{-11}。

式（7-23）表明抗磁磁化率与温度无关。抗磁性来源于电子的轨道运动。因此所有的物质均有抗磁性效应。如惰性气体，大部分多原子气体（如 H_2、N_2），离子型固体（如NaCl），共价键固体（如C、Si、Ge、S和P等），以及大多数有机物质都是抗磁性物质。在顺磁性和铁磁性物质中，抗磁性效应被其他磁性所掩盖了。

7.3.2 顺磁性

若物质的原子次电子层未填满，原子有固有磁矩，在外磁场作用下产生顺磁性。若物质中有自由电子，自由电子即可产生顺磁效应，又可产生抗磁效应，但

顺磁效应大于抗磁效应，总的来说，自由电子产生顺磁效应。

1. 原子（或离子）固有磁矩的顺磁性

若物质的原子有未满壳层的电子，有固有原子磁矩，但是原子受热扰动影响，原子磁矩的方向混乱地分布（铁磁性物质除外），在任何方向都没有净磁矩，对外不显示磁性。而将材料放入外磁场中时，原子磁矩都有沿外磁场方向排列的趋势，感生出和外磁场方向一致的磁化强度。所以磁化率 $\chi > 0$。

在外磁场作用下，原子固有磁矩的磁位能为 $E_H = -\mu_J H\cos\theta$。只有磁位能达到与原子热运动能相当时，即 $kT = H\mu_J$ 时，原子固有磁矩才可能转向外磁场方向。式中，H 为外加磁场的强度；θ 为固有磁矩与外磁场方向的夹角；k 玻耳兹曼常数；T 是热力学温度；μ_J 是原子磁矩。由 $kT = H\mu_J$ 可得出使顺磁性物质原子固有磁矩转向外磁场方向所需要的磁场为 $8\times10^8 \text{A}\cdot\text{m}^{-1}$。在通常的磁场下，原子固有磁矩仅能沿外磁场方向转动一个很小的角度。

假定原子磁矩间彼此是独立的，原子磁矩与外磁场方向的夹角 θ 服从马克斯威—玻耳兹曼（Maxwell- Boltzlman）统计分布规律。朗之万（Langevin）根据这一分布规律导出了在磁场方向的磁化强度 M_A 为

$$M_A = \frac{\mu_0 n \mu_J{}^2 H}{3kT} \tag{7-24}$$

式中，n 为单位体积的原子数；μ_0 是真空磁导率。单位体积物质原子固有磁矩的顺磁磁化率（朗之万顺磁磁化率）χ_A 为

$$\chi_A = \frac{\mu_0 n \mu_J{}^2}{3kT} = \frac{C}{T} \tag{7-25}$$

式中，C 为居里常数。$\chi_A = C/T$ 为居里定律。实践表明只有少数顺磁性物质遵循居里定律，多数顺磁性物质的磁化率遵循居里—外斯定律。

2. 自由电子的顺磁性

某些金属（例如 Cu）的 3d 电子层已填满，4s 电子为自由电子。抗磁性的离子实浸在自由电子气中。如果不考虑离子，仅考虑自由电子，当不存在外磁场时，4s 能带如图 7-4a 所示。当存在外磁场时，正自旋和负自旋电子磁矩的磁位能分别为$-\mu_0\mu_J H$ 和$+\mu_0\mu_J H$，如图 7-4b 所示。正自旋电子能量比负自旋电子能量低 $2\mu_0\mu_J H$，因而有Δn 个自由电子从负自旋的能带迁移到正自旋的能带，以达到图 7-4c 的平衡态。基于这种考虑，泡利（Pauli）得出自由电子磁化率为

$$\chi_{e(\text{p})} = \frac{\mu_0 m \mu_B{}^2 (3n)^{1/3}}{\hbar^2 \pi^{4/3}} \tag{7-26}$$

式中，n 是单位体积的自由电子数；$\hbar = h/(2\pi)$，h 为普朗克常数；m 为电子的静止质量。式（7-26）表明自由电子磁化率与温度无关。后来 Landau 证明，在磁场下，

自由电子也受洛伦兹力的作用而产生附加的抗磁性效应，结果使自由电子的顺磁磁化率 χ 比式（7-26）的小 1/3。

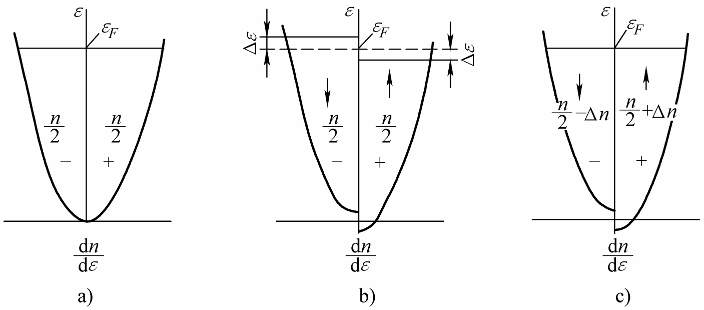


图 7-4　自由电子的泡利顺磁性

a）$H=0$　b）外磁场作用下的磁位能　c）自由电子的迁移

总之，非铁磁性物质是抗磁性或是顺磁性，取决于抗磁性和顺磁性这两项中哪一项大（顺磁性项包括原子固有磁矩的顺磁性和自由电子的顺磁性）。

7.3.3　抗磁性金属与顺磁性金属

在磁场的作用下电子的循轨运动要产生抗磁矩，而离子的固有磁矩则产生顺磁矩。此外，还要看到，自由电子在磁场的作用下也产生抗磁矩和顺磁矩，不过它所产生的抗磁矩远小于顺磁矩，故自由电子的主要贡献是顺磁性。金属均由离子和自由电子所构成，因此对于一种金属来说，其内部既存在着产生抗磁性的因素，又存在着产生顺磁性的因素，属于哪种磁性金属，取决于哪种因素占主导地位。

金属的离子，由于核外电子层结构不同，可以分为两种情况：首先是它的电子壳层已全部被填满，即固有磁矩为零。在外加磁场的作用下由核外电子的循轨运动产生抗磁矩，抗磁矩的强弱取决于核外电子的数量。如果离子部分总的抗磁矩大于自由电子的顺磁矩，则金属为抗磁金属。属于这种情况的抗磁金属有铜、金和银等。锑、铋和铅等金属也属于这种情况，所不同的是它们的自由电子向共价键过渡，因而呈现出异常大的抗磁性。还有些金属，如碱金属和碱土金属，它们的离子也是填满的电子结构，但它们的自由电子所产生的顺磁性大于离子部分的抗磁性。因此它们是顺磁性金属，如铝、镁、锂、钠和钾等。其次是离子有未被填满的电子层，即离子具有较强的固有磁矩。在外磁场的作用下，这些固有磁矩所产生的顺磁矩远大于核外电子循轨运动所产生的抗磁矩。具有这种离子的金属都有较强的顺磁性，它们属于强顺磁性金属。如 3d−金属中的钛和钒等；4d−金属中的铌、锆和钼等；5d−金属中的铪、钽、钨和铂等。多数金属属于顺磁性。

7.4　铁磁体（包括铁氧体）自发磁化

有些物质在放入外磁场中时，磁化率 $\chi > 0$，且其数值很大，约 10^1～10^6。实验发现，这些物质的磁化曲线 M～H 是非线性的复杂函数，反复磁化时出现磁滞现象，这种物质称铁磁性物质。铁磁性物质的原子不仅有固有原子磁矩，而且原子磁矩分区间（该区间称磁畴）地自发平行取向，所以原子磁矩非常容易朝外磁场方向排列，只要在很小的外磁场强度 H 下就可以感生出很大的磁化强度 M。但是当温度高于某个临界值 T_c（称居里温度）时，材料的铁磁性将转变成顺磁性，这时磁化率 χ 服从居里—外斯定律。下面主要讨论自发磁化的本质。

7.4.1　外斯（P. E. Weiss）铁磁性假说

如上所述，铁磁性材料的磁化率远大于顺磁性物质，就使铁磁性材料具有十分有用的磁特性，如很高的导磁率、有剩磁、矫顽力等。为了说明铁磁性物质比顺磁性或抗磁性物质有强得多的磁性，1907 年法国物理学家外斯（P. E. Weiss）在朗之万顺磁理论的基础上首先系统地提出了铁磁性假说（简称分子场理论），其主要内容为：①铁磁物质内部存在很强的“分子场”，它使原子磁矩同向平行排列，即自发磁化到饱和；②铁磁体的自发磁化分成若干磁畴（自发磁化到饱和的区域称为磁畴），由于磁体中各磁畴的磁化方向不一致，所以大块磁体对外不显示磁性。

设分子场 H_m 使铁磁体自发磁化，铁磁体的自发磁化强度 $M_s(T)$与分子场 H_m 成正比，即

$$H_m = \lambda M_s(T) \qquad (7\text{-}27)$$

式中，λ 为分子场系数。在温度大于 0K 时，由于原子的热振动，分子场仅能使原子磁矩在一定程度上平行排列。设单位体积中磁性原子数为 n，在分子场和外磁场作用下铁磁体的宏观磁化强度 M 随温度 T 和外磁场 H 的变化用玻耳兹曼统计可得

$$M = M(0)B_J(y) \qquad (7\text{-}28)$$

式中，$M(0) = nJg_J\mu_B$；$B_J(y)$为布里渊函数，为

$$B_J(y) = \left[\frac{2J+1}{2J}\coth\left(\frac{2J+1}{2J}y\right) - \frac{1}{2J}\coth\frac{y}{2J}\right] \qquad (7\text{-}29)$$

式中，$y = Jg_J\mu_B\mu_0(H+\lambda M_s)/(kT)$。解式（7-28）和式（7-29）可以求出在一定的磁场和温度下的磁化强度。当外加磁场为零时，可以求出铁磁性物质的自发磁化强度 $M_s(T)$。解上述公式得出以下三个主要结论：①在 $T < T_c$ 的任何温度下，自发磁化

总是存在的，材料表现出铁磁性。当 $T>0K$ 时，$B_J(y)>1$，$M_s(T>0K)=nJg_J\mu_B$，温度升高，自发磁化强度逐渐降低。在 $T\geqslant T_c$ 后，如果外加磁场为零，式（7-28）和式（7-29）没有共同的解，即自发磁化强度为零，材料表现出顺磁性。这个临界温度就是居里温度 T_c；②材料的磁化率 χ 和温度 T 服从居里—外斯定律，即

$$\chi=\frac{C}{T-\Theta} \tag{7-30}$$

式中，C 是居里常数。对于铁磁体 Θ 为正，等于居里点 T_c。当 $T=\Theta$ 时，铁磁性转变为顺磁性；当 $T>\Theta$ 时，铁磁性物质磁化率服从居里—外斯定律。对于反铁磁性物质 Θ 取负值，当 $T>T_N$（反铁磁性物质的奈耳点）时，其磁化率也服从居里—外斯定律。这些结果与实验结果符合得很好。图 7-5 为铁磁性与反铁磁性物质 $\chi-T$ 关系；③分子场系数 λ 与居里温度 T_c 成正比，即

$$T_c=\frac{\mu_0 nJ(J+1)g_J{}^2\mu_B{}^2}{3k}\lambda \tag{7-31}$$

后来的量子自发磁化理论证明这一结果是正确的。

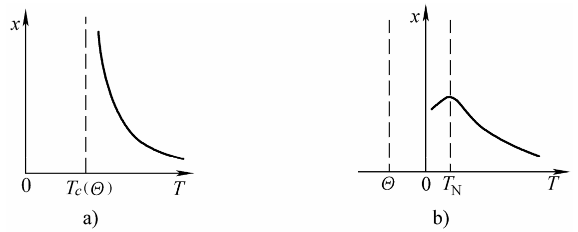


图 7-5 铁磁性与反铁磁性物质 $\chi-T$ 关系

a）铁磁性物质 b）反铁磁性物质

以上结果表明，外斯的分子场理论大体上能描述铁磁体的自发磁化。但是在外斯分子场理论中，最难以解释的是铁磁体为什么有的自发磁化？外斯分子场理论仅是一种唯象理论，并没有说明分子场的本质。

7.4.2 直接交换作用

根据式（7-28）和式（7-31）估算一下分子场的大小。对于铁来说，$T_c=1043K$，$n=8.49\times10^{28}m^{-3}$。设 $J(J+1)g_J\mu_B\approx n_{eff}\mu_B=2.2\times9.273\times10^{-24}=2.04\times10^{-23}A\cdot m^2$（$n_{eff}$ 为原子磁矩的实验值），代入式（7-31）得 $\lambda=9.73\times10^2$，铁的 $M_s=n_{eff}n\mu_B\approx1.73\times10^6A\cdot m^{-1}$，代入式（7-28）得分子场约 $1.73\times10^9A\cdot m^{-1}$。这一磁场比晶体中一个原子磁矩在相邻原子处产生的磁场 $8\times10^5A\cdot m^{-1}$ 大 3～4 个数量级，说明分子场不起源于磁的相互作用。

当 1924～1926 年建立了描述微观电子运动的量子力学理论后，1928 年海森堡

（W. Heisenberg）和弗伦克尔（Френкелъ）几乎同时分别提出分子场是相邻原子间电子自旋的交换作用的理论。这种交换作用是一种量子力学效应，是属于静电性质的。下面用氢分子模型说明交换作用原理。

图 7-6 中 a、b 分别代表两个氢原子核，1 和 2 代表两个电子。在氢分子内原子核之间，原子核与电子间的距离分别如图 7-6 所示。当两个氢原子相距很远时，两原子的电子云不重叠，它们有各自能量 E_0，两个孤立的氢原子的能量和为 $2E_0$，同时两个氢原子间的电子自旋运动也相互不影响。当两个氢原子靠近到能形成氢分子时，两个氢原子就要结合在一起。这时电子的运动与两个孤立氢原子时的运动就不同了，相应地整个氢分子的能量与两个孤立原子的能量和也不同了。原子磁矩间磁的相互作用很弱不予考虑，而电的作用有如下几项：原子核之间的静电相互作用；1 和 2 两个电子间的电相互作用；a 核与电子 1、2 之间的电相互作用；b 核与电子 1、2 之间的电相互作用。两孤立的氢原子结合成分子后，其能量为 $E=2E_0+E'$。根据量子力学的理论计算，E' 不仅决定于上述的静电作用，还与两个电子间自旋的相对取向有关。由量子力学的理论计算得到氢分子的能量为

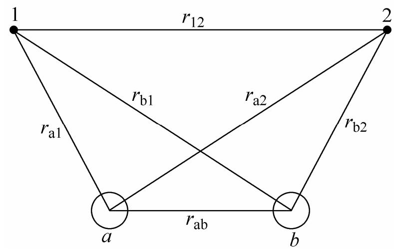


图 7-6　氢分子模型

$$E=2E_0+\frac{e^2}{r_{ab}}+K-\frac{A}{2}-2A\boldsymbol{\sigma}_1\boldsymbol{\sigma}_2 \tag{7-32}$$

式中，K 和 A 为

$$\left.\begin{aligned}K&=\iint\psi_a^*(1)\psi_b^*(2)\left(\frac{e^2}{r_{12}}-\frac{e^2}{r_{a2}}-\frac{e^2}{r_{b1}}\right)\psi_a(1)\psi_b(2)\mathrm{d}\tau_1\mathrm{d}\tau_2\\A&==\iint\psi_a^*(2)\psi_b^*(1)\left(\frac{e^2}{r_{12}}-\frac{e^2}{r_{a1}}-\frac{e^2}{r_{b2}}\right)\psi_a(2)\psi_b(1)\mathrm{d}\tau_1\mathrm{d}\tau_2\end{aligned}\right\} \tag{7-33}$$

式中，$\psi_a(1)$和 $\psi_a(2)$分别代表电子 1 和电子 2 在 a 核周围运动的波函数。$\psi_a^*(1)$和 $\psi_a^*(2)$分别是它们对应波函数的共轭波函数。$\psi_b(1)$和 $\psi_b(2)$分别代表电子 1 和电子 2 在 b 核周围运动的波函数。$\psi_b^*(1)$和 $\psi_b^*(2)$分别是它们相应的共轭波函数。$\mathrm{d}\tau_1$ 和 $\mathrm{d}\tau_2$ 分别代表空间的体积元。K 是库仑相互作用的平均势能，它与两个电子之间的自旋取向无关。A 是交换积分常数。在 A 项内包括了 $\psi_a(2)$和 $\psi_b(1)$的波函数，它代表电子 1 和电子 2 相互交换位置时对应的能量。A 的大小与两电子云重叠的程度有关。

$-2A\boldsymbol{\sigma}_1\boldsymbol{\sigma}_2$ 项是与两个电子自旋取向有关的能量，称为自旋交换能，用 E_{ex} 表示。式中 $\boldsymbol{\sigma}_1$ 和 $\boldsymbol{\sigma}_2$ 是两电子以 $\hbar$ 为单位的电子自旋角动量，即 $P_s/\hbar=\sqrt{s(s+1)}=\sigma$，$P_s$ 为

自旋角动量。感兴趣的是两个电子自旋相对取向对能量的影响，因此把式（7-32）中的前四项能量看作是常数，而仅研究最后一项能量，即

$$E_{ex} = -2A\boldsymbol{\sigma}_1\boldsymbol{\sigma}_2 \tag{7-34}$$

为简便起见，把式（7-34）按经典矢量模型来处理，设两个电子自旋角动量矢量 $\boldsymbol{\sigma}_1$ 和 $\boldsymbol{\sigma}_2$ 的夹角为 φ，这样可把式（7-34）写成

$$E_{ex} = -2A\sigma_1\sigma_2\cos\varphi \tag{7-35}$$

可见当交换积分 A 为正时，交换能为最小值的条件是相邻原子间的电子自旋角动量同向平行排列（$\varphi = 0°$，$\cos\varphi = 1$）。当交换积分 A 为负时，交换能最小的条件是相邻原子的电子自旋角动量反向平行排列（$\varphi = 180°$，$\cos\varphi = -1$）。

根据能量最小值的原理可得，当铁磁物质内部相邻原子的电子交换积分 A 取正值时，相邻原子自旋磁矩要同向平行排列，从而实现自发磁化到饱和。这就是铁磁性的起因。

原子磁矩不为零，交换积分常数 $A > 0$ 是铁磁性的必要与充分条件。由式（7-33）可知，相邻原子的电子云（波函数）在重叠区有相同符号，所以 $\psi_a^*(1)$、$\psi_a^*(2)$和 $\psi_a(2)$、$\psi_b(1)$都是正的，要想使 $A > 0$，必须满足

$$\frac{1}{r_{12}} > \frac{1}{r_{a1}} + \frac{1}{r_{a2}} \tag{7-36}$$

可见要使 $A > 0$，必须要求 r_{12} 要足够小，即要求未填满的次电子层的电子云在两原子核之间重叠。不过重叠区也不能过分大。3d 金属 Fe、Co 和 Ni 及它们的合金正好符合这一条件。因为 3d 电子的轨道角量子数 $l = 2$，其轨道半径较小，两电子云既重叠，但重叠区又不过分大，保证了 $A > 0$，因此它们是铁磁性的。根据这些条件，斯莱特（Slater）总结出交换积分常数 A 和 a/r_{3d} 的关系，如图 7-7 所示。a 是相邻原子间距，r_{3d} 是未填满电子的 3d 电子层的半径。当 $a/r_{3d} > 3$ 时，$A > 0$，为铁磁性。但 a 也不能过大。若 a 过大，3d 电子云不重叠，交换作用很弱，使 $A \to 0$，为顺磁性。若 a 过小，就会使得式（7-36）不成立，从而使得 $A < 0$，导致反铁磁性。在 Mn 和 Cr 中通过添加第二或第三组元，扩大 a 值则有可能使反铁磁性转变为铁磁性。

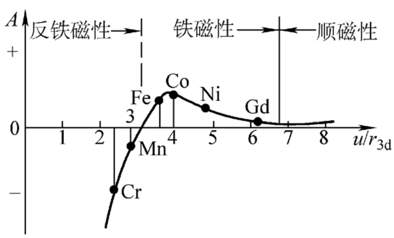


图 7-7　交换积分 A 和 a/r_{3d} 的关系

7.4.3　间接交换作用

间接交换作用通常在氧化物（如铁氧体）和稀土金属及其合金中存在。不过

铁氧体和稀土金属中的间接交换作用又不相同。

1. 稀土金属化合物中的间接交换作用

稀土金属中对磁性有贡献的 4f 电子是局域化的。4f 电子层半径为 0.05～0.06nm，外层的 $5p^6 5d^1 6s^2$ 电子对 4f 电子起屏蔽作用。相邻原子的 4f 电子云不重叠，不可能存在直接交换作用。茹德曼（Ruderman）、基特尔（Kittel）、胜谷（Kasuya）和良田（Yosida）等人提出了导电电子与内层电子的交换作用理论，称为 RKKY 理论。RKKY 理论中心思想是：在稀土金属中 4f 电子是局域的，6s 电子是游动的。f 电子与 s 电子发生交换作用，使 s 电子极化，这个极化了的 s 电子的自旋对 f 电子自旋取向有影响，结果形成了以游动的 s 电子为媒介，使磁性离子的 4f 电子自旋与相邻的离子的 4f 电子自旋存在间接交换作用，从而产生自发磁化。

在稀土金属间化合物中，例如富 3d 过渡族化合物，如 ReM_5、Re_2M_5 等已成为重要的永磁材料。这类化合物的晶体结构都是由 $CaCu_5$ 型六方结构派生出来的。其中 ReM_5，如 $ReCo_5$ 的结构与 $CaCu_5$ 的结构相同，如图 7-8 所示。在这类化合物中 R-R 和 R-M 原子间距都较远，无论是 4f 电子之间还是 3d—4f 电子云之间都不可能有重叠。它也是以传导电子为媒介产生间接交换作用，对于原子序数小于 Gd 的轻稀土金属来说，3d 金属原子与 4f 稀土金属原子磁矩平行排列，导致两种原子磁矩铁磁性耦合；而原子序数大于 Gd 的重稀土金属，3d 金属原子与 4f 稀土金属原子磁矩的反平行排列，导致了两种原子磁矩的亚铁磁性耦合，如图 7-9 所示。

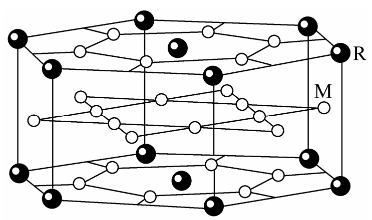


图 7-8　ReM_5 的晶体结构

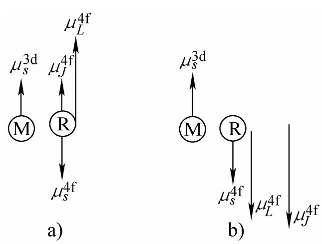


图 7-9　稀土化合物中原子磁矩的耦合方式

a）轻稀土化合物　b）重稀土化合物

2. 铁氧体中的间接交换作用

在氧化铁中，Fe-Fe 原子间的距离是 0.428nm。由于铁原子间距过大，3d 电子间不可能有直接交换作用。为了解释氧化物中的自发磁化，1934 年克拉默尔斯（Kramers）提出了间接交换作用（又称为超交换作用）。到了 1950 年安德森（Anderson）进一步完善了超交换作用理论。下面定性地作一简介。

图 7-10 是 MnO 的晶胞。图中两个斜影线画出的对角面把它分成两个磁矩相反的次晶格。它的特点是 Mn^{2+}和 O^{2-}交替地占据晶格的位置。任何一个 Mn^{2+}的最近邻都是 O^{2-}，而每一个 O^{2-}的周围又都是以 Mn^{2+}作最近邻。每一个 O^{2-}两侧（上

下，左右，前后）相距为 a 的 Mn^{2+}的磁矩都是反平行的。另一种情况是通过 O^{2-}形成 90°，间距为$(\sqrt{2}/2)a$ 的 Mn^{2+}的磁矩可能是反向的，也可能是同向的。例如图 7-10 中 1 和 2 同向，2 和 3 反向。显然决定离子磁矩相对取向的不是 Mn^{2+}和 Mn^{2+}间的直接交换作用，而是通过 O^{2-}所产生的一种间接交换作用。下面分析一下 Mn^{2+}和 O^{2-}的相互作用。

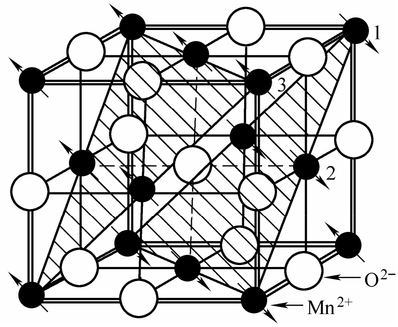


图 7-10　MnO 晶胞中 Mn^{2+}和 O^{2-}的分布及离子磁矩方向

一对 Mn^{2+}、O^{2-}离子处于基态时，Mn^{2+}离子 $3d^5$ 亚层的 5 个电子自旋方向相同，而 O^{2-}离子 $2p^6$ 亚层的 6 个电子自旋方向两两相反。O^{2-}离子的外层电子是满的，它不可能与周围的 Mn^{2+}的 3d 亚层产生交换作用。但由于 Mn^{2+}的 3d 亚层没有填满电子，因而 O^{2-}离子的 2p 电子就有可能跑到 Mn^{2+}的 3d 层中的空穴中去。当 O^{2-}的一个电子一旦跑到 Mn^{2+}的 3d 层中去时，这个 2p 态电子就变成了 3d 态电子。因而系统就成了激发态。根据泡利原理，这个激发态的电子自旋磁矩就要与 Mn^{2+}的磁矩方向相反。O^{2-}失去一个 2p 电子后，就变成了 O^-，而成为一个具有一个电子自旋磁矩的磁性离子。O^-的磁矩方向仍然和激发到 3d 态电子的自旋磁矩的方向反向平行，这是因为考虑的激发态是一级激发态，一级激发态的电子只改变位置，而不改变方向。这时 Mn^+的 $3d^6$ 的电子自旋方向为↑↓↑↑↑↑，而 O^-的 $2p^5$ 的电子自旋方向为↑↓↑↓↑。这样就确定了 Mn^+和 O^-的磁矩方向相同。而 O^-是磁性离子，它与邻近的别的磁性离子 Mn^{2+}发生交换作用，从而确定它与这些近邻的 Mn^{2+}磁矩的相对取向，再间接地确定各 Mn^{2+}之间的磁矩相对取向。

因为 p 态电子云的角分布是“哑铃”状的。因此 O^{2-}的电子云只与其两侧成一直线的两个 Mn^{2+}的电子云有重叠，如图 7-11 所示。因此 O^{2-}的 2p 电子只跑到两侧的 Mn^{2+}中去，不会跑到垂直于“哑铃”方向的 Mn^{2+}中去。当 O^{2-}中的一个 2p 电子跑到左边的 Mn^{2+}的 3d 层中去时，O^-就要和右边的 Mn^{2+}产生较强的交换作用，而在垂直于“哑铃”方向的交换作用很弱。这是因为 O^-和它右边的 Mn^{2+}的电子云重叠较多，而和垂直于这个“哑铃”方向的邻近的 Mn^{2+}的电子云几乎没有重叠。“哑铃”状分布的电子是磁量子数 m_l 相同而自旋方向相反的两个电子。计算证明 O^-与右边 Mn^{2+}的交换积分 A 是负的，因而和右边的 Mn^{2+}的磁矩方向必定反平行。Mn^+ —O^-—Mn^{2+}的磁矩方向如图 7-12 所示。因此在 O^{2-}两侧成一直线上的两个 Mn^{2+}的磁矩必然是反平行的。称这种通过氧离子而确定锰离子磁矩相对取向的交换作用为间接交换作用或超交换作用。这种交换作用就使得 MnO 中的 Mn^{2+}磁矩，一半向着一个方向，另一半向着相反的方向，而总的磁矩为零。固此 MnO 是反铁

磁性的。

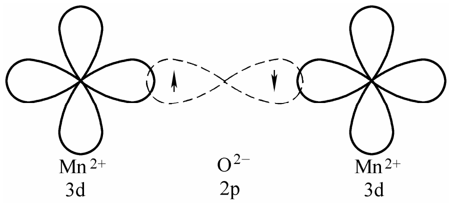


图 7-11　Mn^{2+}–O^{2-}–Mn^{2+}电子云角分布

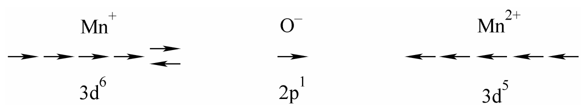


图 7-12　Mn^{+}–O^{-}–Mn^{2+}的磁矩方向

铁氧体有三种晶体结构类型，即尖晶石型、石榴石型和磁铅石型。前者是最常见的一种。图 7-13 是铝镁尖晶石（$M_xO{\cdot}Al_2O_3$，其中 M 代表二价金属离子，如 Mn^{2+}、Co^{2+}、Mg^{2+}、Ni^{2+}、Zn^{2+}等）晶体的一个晶胞。可见直径大的 O^{2-}（0.265nm）占据了晶胞中大部分空间，直径比较小的金属离子 Mg^{2+}（0.156nm）和 Al^{3+}（0.116nm）嵌镶在 O^{2-}离子之间的空隙里，而且有一定的规律。在 Mg^{2+}的四周有四个最近邻的 O^{2-}，如果把这四个 O^{2-}的中心联接起来，就构成了一个正四面体。Mg^{2+}处在这四面体的中心位置。而 Al^{3+}最近邻有六个 O^{2-}，这六个 O^{2-}的中心联线构成正八面体。Al^{3+}处于正八面体的中心位置，如图 7-14 所示。分析晶胞中所有离子的相对位置，就会发现 O^{2-}之间的间隙只有这两种。称四面体中心位置为 *A* 位置，八面体的中心为 *B* 位置。占据 *A* 位置的金属离子所构成的晶格为 *A* 次晶格。占据 *B* 位置的金属离子所构成的晶格为 *B* 次晶格。

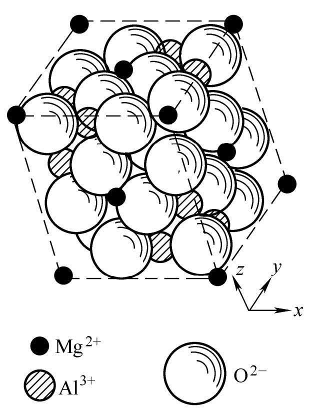


图 7-13　尖晶石型晶体结构

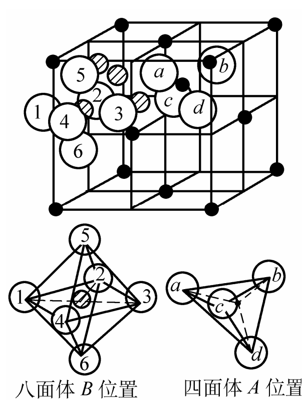


图 7-14　四面体和八面体示意图

A 位置或 B 位置的金属离子间都要通过 O^{2-}发生间接交换作用。A 和 B 位置上离子的磁矩是反铁磁耦合，在铁氧体中往往是 A、B 两个位置上的磁矩不等，因而出现了亚铁磁性。例如单尖晶石铁氧体的结构式为

$$\underset{(A\text{位置})}{(M_{1-x}^{2+}Fe_{x}^{3+})}\underset{(B\text{位置})}{(M_{x}^{2+}Fe_{2-x}^{3+})}O_4$$

A 位置上的磁矩和为$[5x+m_x(1-x)]\mu_B$，式中，x 为原子数；$m_x\mu_B$ 为金属离子 M^{2+}的磁矩；$5\mu_B$ 为 Fe^{3+}的磁矩。B 位置上的总磁矩为$[m_xx+5(2-x)]\ \mu_B$。面分子的磁矩 m 为 A 位与 B 位上磁矩之差，即

$$\begin{aligned} m &= [m_x x + 5(2-x)]\mu_B - [5x - m_x(1-x)]\mu_B \\ &= 10(1-x)\mu_B + m_x(2x-1)\mu_B \end{aligned} \tag{7-37}$$

$MnFe_2O_4$ 和 $NiFe_2O_4$ 分子磁矩的计算值与实验值基本相符合。

在实际使用中单铁氧体在磁性上不能满足要求，于是人们就根据实际需要将两种或两种以上的单铁氧体按一定的比例制备成多元系铁氧体，称为复合铁氧体。

铁氧体磁性材料的特点是电阻率特别高（比金属磁性材料的电阻率高 100 万倍），这在高频和超高频技术中应用，就有很大的优越性。另外，铁氧体的原材料来源丰富，成本低。但铁氧体的居里点偏低，温度稳定性差，M_s 也较低，故在低频和高功率下一般还是用金属磁性材料。

7.5 铁磁体中的磁晶各向异性、磁致伸缩

交换作用能使铁磁物质中的相邻原子磁矩同向平行（铁磁性耦合）或反向平行（反铁磁性耦合）排列，在磁畴范围内使原子磁矩自发磁化到饱和，但不可能使整个大块的铁磁体自发磁化到饱和。因大块铁磁体磁化到饱和后，退磁能要大大的提高，它迫使铁磁体分成畴。磁畴的大小、形状、取向与铁磁体的磁晶各向异性能、退磁场能、磁弹性能、交换能等有关。这些能量对铁磁体的磁行为和磁参量有重要的影响。在这些能量中，交换能是近程的，属于静电性质的，其数值比其他各项能量大 3～4 个数量级。其他各项能量属于静磁相互作用性质。下面介绍铁磁体中的磁晶各向异性能等各项能量。

7.5.1 磁晶各向异性能

在单晶体的不同晶体学方向上，其光学、电学、热膨胀、力学和磁学性能都不同。这种特性称为晶体的各向异性。单晶体的磁性各向异性称为磁晶各向异性。例如铁、镍、钴的单晶体各个晶向的磁化曲线是不同的，如图 7-15 所示。与求固体的弹性变形功相似，铁磁体在磁化时，外磁场对铁磁体所做的功 W 为

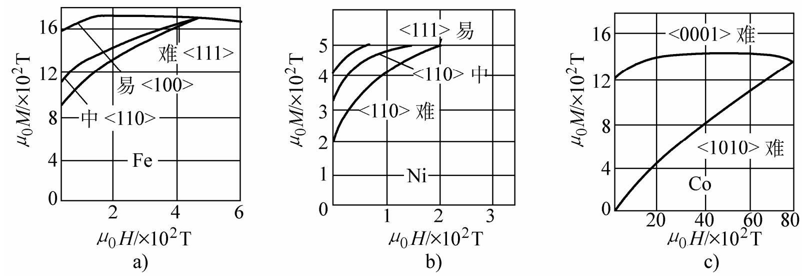


图 7-15　铁、镍、钴沿不同晶向的磁化曲线

a）铁　b）镍　c）钴

$$W = \int_0^{M_s} \mu_0 H \mathrm{d}M \qquad (7\text{-}38)$$

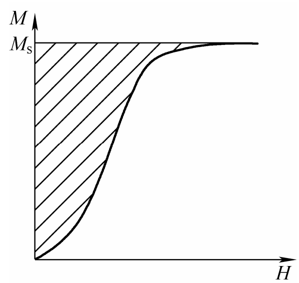


图 7-16　磁化功示意图

式中，μ_0 是真空磁导率；H 为外磁场强度；M_s 为饱和磁化强度。W 相当于图 7-16 的阴影面积。因此体心立方的铁<100>方向为易磁化方向，而<111>方向为难磁化方向。对于镍单晶体与铁正好相反，它是面心立方结构，易磁化方向为<111>，<100>方向为难磁化方向。钴是六角晶体，易磁化方向是<0001>方向，即与六边形柱体轴的方向重合。

立方晶体沿<uvw>方向磁化和沿<001>方向磁化功的差 $E_K = W_{\langle uvw\rangle} - W_{\langle 001\rangle}$称为磁晶各向异性能。很明显，磁晶各向异性能是磁化方向或磁化强度方向的函数。对立方晶体，设 α_1、α_2 和 α_3 分别是磁化强度与三个晶轴的方向余弦。所以 $E_K = f(\alpha_1, \alpha_2, \alpha_3)$。这是多元函数，将它按泰勒级数展开，并根据晶体的对称性和三角函数的关系式，可得

$$E_K = K_1(\alpha_1^2\alpha_2^2 + \alpha_2^2\alpha_3^2 + \alpha_3^2\alpha_1^2) + K_2(\alpha_1^2\alpha_2^2\alpha_3^2) \qquad (7\text{-}39)$$

式中，K_1、K_2 称为磁晶各向异性常数。当 K_2 很小时，可以只用 K_1 来描述立方晶体的磁晶各向异性能 E_K。对铁来说，K_1 是正的，对镍来说，K_1 是负的。

对其他结构晶体，磁晶各向异性能也与其磁晶各向异性常数有关，六方晶体对称性差，各向异性大。

磁晶各向异性常数 K_1 和 K_2 或 K_1+K_2 是衡量材料的磁各向异性大小的重要常数，它的大小与晶体的对称性有关。晶体的对称性越低，它的 K_1+K_2 的数值越大。K_1 和 K_2 是内禀特性。也即主要决定材料成分。

磁晶各向异性的起源有多种解释。用自旋—轨道相互作用解释的中心思想是，磁晶各向异性和晶体场对电子轨道运动的影响有关。一方面电子轨道磁矩产生的

磁场对电子自旋运动作用，使轨道和自旋间存在耦合作用；另一方面电子轨道平面受晶体场作用能量间并被消除，这两方面的作用叠加在一起，就使得原子磁矩倾向于在晶体的某些方向上能量低，而在另一些方向上能量高。原子磁矩能量低的方向为易磁化方向，而能量高的方向为难磁化方向。在无外磁场作用的平衡状态下，原子磁矩倾向于排列在易磁化方向上。

7.5.2 退磁场能

铁磁体与外磁场的相互作用能称为静磁能（或磁位能）。铁磁体与自身退磁场的相互作用能称为退磁场能。实验表明，磁性材料被磁化后，只要材料不是闭合形状或者无限长，材料内部的总磁场 H 将小于外磁场 H_e，这是因为非闭合的磁性材料被磁化后在其端面将会有正负磁荷出现。这些磁荷将在材料内外产生一个退磁场 H_d，H_d 的方向在材料内部与 H_e 和磁化强度 M 方向相反，其作用是削弱外磁场。退磁场越大，材料磁化越不容易，退磁场的表达式为

$$H_d = -NM \tag{7-40}$$

式中，负号表示 H_d 与 M 的方向相反；N 为退磁因子，其大小与铁磁体的形状有关。球状铁磁体在 x、y、z 三坐标轴上（坐标原点在球心）的退磁因子 N_x、N_y、N_z 的关系为 $_x+N_y+N_z = 1$，$N_x = N_y = N_z = 1/3$。有限长的长旋转椭球体沿长轴 l 方向和短轴 d 方向的退磁因子 N_l、N_d 为

$$\left.\begin{aligned} N_l &= \frac{1}{k^2-1}\left(\frac{k}{2\sqrt{k^2-1}}\ln\frac{k+\sqrt{k^2-1}}{k-\sqrt{k^2-1}}-1\right) \\ N_l &+ 2N_d = 1 \end{aligned}\right\} \tag{7-41}$$

式中，$k = l/d$。随铁磁体长度的增加，退磁因子减少，即退磁场减少。

如果材料不是均匀磁化，则退磁因子不仅和尺寸有关，还和材料的磁导率有关。

铁磁性材料与自身退磁场的相互作用能 E_d（称为退磁场能）可以写成

$$E_d = \int_0^M \mu_0 H_d \mathrm{d}M = \frac{\mu_0 NM}{2} \tag{7-42}$$

7.5.3 磁致伸缩

在磁场中磁化时，铁磁体的尺寸或体积发生变化的现象称为磁致伸缩。通常用磁致伸缩系数 $\lambda = \Delta l/l$ 来描述铁磁体尺寸大小的相对变化（l 为铁磁体的长度）。磁致伸缩系数随磁场的增强而增加，当磁场达到一定数值后，它达到饱和值，称为饱和磁致伸缩系数 λ_s。λ_s 有正值也有负值。根据单晶体的各向异性和对称性可以得出立方晶体的饱和磁致伸缩系数的表达式为

$$\lambda_s = \frac{3}{2}\lambda_{110}\left(\alpha_1{}^2\beta_1{}^2 + \alpha_2{}^2\beta_2{}^2 + \alpha_3{}^2\beta_3{}^2 - \frac{1}{3}\right) + 3\lambda_{111}\left(\alpha_1\alpha_2\beta_1\beta_2 + \alpha_2\alpha_3\beta_2\beta_3 + \alpha_3\alpha_1\beta_3\beta_1\right) \tag{7-43}$$

式中，α_i 和 β_i（i =1, 2, 3）分别是磁化强度和测量方向与立方晶体的三个晶轴夹角的方向余弦；λ_{100} 和 λ_{111} 分别是沿<100>和<111>晶轴方向磁化到饱和时的饱和磁致伸缩系数。

当晶体的磁致伸缩是各向同性时，则 $\lambda_{100}=\lambda_{111}=\lambda_0$，式（7-43）变成

$$\lambda_s = \frac{3}{2}\lambda_0\left(\alpha_1\beta_1 + \alpha_2\beta_2 + \alpha_3\beta_3 - \frac{1}{3}\right) = \lambda_0\frac{3}{2}\left(\cos^2\theta - \frac{1}{3}\right) \tag{7-44}$$

式中，θ 是磁化强度方向和测量磁致伸缩方向之间的夹角。例如当 $\theta = 0$，$\lambda_s = \lambda_0$；又 $\theta = \pi/2$，$\lambda_s = -\lambda_0/2$，说明当纵向伸长时，横向要收缩。

立方晶体的多晶体磁致伸缩系数 λ 与单晶体的磁致伸缩系数 λ_{100} 和 λ_{111} 的关系为

$$\lambda = \frac{2}{5}\lambda_{100} + \frac{3}{5}\lambda_{111} \tag{7-45}$$

对于 3d 金属及合金，λ_s 为相当于温度变化 1℃时，由热膨胀所引起的线性变化。但是一些大磁致伸缩材料，例如 $Tb_{0.27}Dy_{0.73}Fe_{1.9}$，它的磁致伸缩系数可达到 2500×10^{-6}，被用于制动器和声纳之中。磁致伸缩现象对铁磁体的畴结构、技术磁化行为及某些技术磁参量也有重要的影响。

当材料中存在内应力或外加应力时，磁致伸缩和应力相互作用，与此有关的能量称为磁弹性能 E_σ。在立方晶系各向同性材料中，磁弹性能 E_σ 为

$$E_\sigma = \frac{3}{2}\lambda_s\sigma\sin^2\theta \tag{7-46}$$

式中，σ 是应力；θ 是磁化方向和应力方向的夹角。当 λ_s 和 σ 符号相同，$\theta = 0$ 时，磁弹性能最小，应力的方向是易磁化方向。而 $\theta = \pi/2$ 时，磁弹性能最大，在垂直应力的方向是难磁化方向；当 λ_s 和 σ 符号相反时，$\theta = 0$ 时能量最大，沿应力的方向是难磁化方向。而 $\theta = \pi/2$ 的方向磁弹性能最小，垂直应力的方向应是易磁化方向。由于应力使铁磁体变成各向异性时，称为应力各向异性。当磁晶各向异性很小时，应力各向异性就有重要作用。

7.6　畴壁与磁畴结构

实验已证明，在铁磁体内确实存在磁畴。图 7-17 是 Si-Fe 合金在（001）晶面上观察到的磁畴，它由

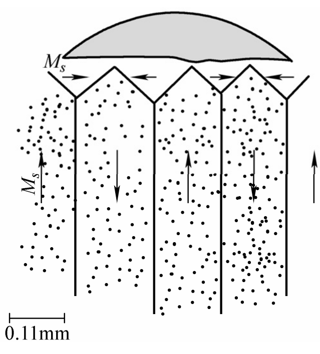


图 7-17　Si-Fe 合金在（001）晶面上的片状畴和封闭畴

片状畴和三角畴组成。畴与畴之间的边界称为畴壁。相邻两个片状畴的磁矩夹角为 180°时，它们的边界称为 180°畴壁。片状畴与三角畴（又称封闭畴）之间磁矩相互垂直，它们的边界称为 90°畴壁。在每一个磁畴内，磁矩平行或反平行地有序排列，产生自发磁化。而不同磁畴的自发磁化矢量则是随机排列。磁畴的形状、尺寸、磁畴壁的厚度由交换能、退磁场能、磁晶各向异性能及磁弹性能来决定。平衡状态的磁畴结构和磁畴壁应具有最小的能量。

7.6.1 畴壁

磁畴和磁畴之间的边界称为畴壁。在畴壁内原子磁矩的方向逐渐转变。根据原子磁矩转变的方式,可将畴壁分为布洛赫壁（Bloch Walls）和奈尔壁（Neel Wall），如图 7-18 所示。布洛赫壁的特点是畴壁内的磁矩方向改变时始终与畴壁平面平行，一般在大块的铁磁性材料内存在布洛赫壁。当铁磁体厚度减少到相当于二维的情况，即厚度为 1~10nm 的薄膜时，则畴壁的磁矩始终与薄膜表面平行地转变，这种畴壁称奈尔壁。由于畴壁内部的原子磁矩不再相互平行，磁矩间的交换作用能就有所提高，同时，由于在畴内磁矩偏离了易磁化方向，磁各向异性能也有所提高。因此，和磁畴内比，畴壁是高能区域。

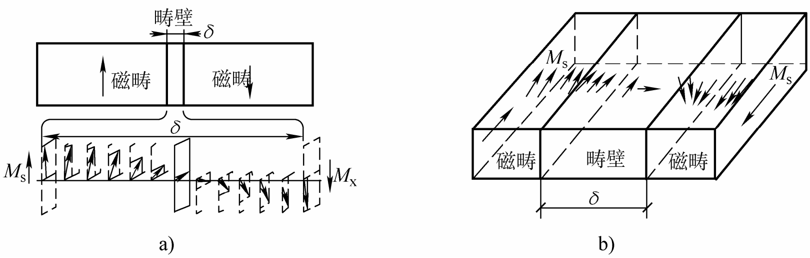


图 7-18 布洛赫壁和奈尔壁中磁矩过渡的方式

a）布洛赫壁中磁矩过渡的方式 b）奈尔壁中磁矩过渡的方式

1932 年布洛赫首先从能量的观点分析了 180°布洛赫壁。如果原子磁矩在相邻两原子间突然反向，则交换作用能的改变 ΔE_{ex} 为

$$\Delta E_{ex} = E_{ex}[\uparrow\downarrow] - E_{ex}[\uparrow\uparrow] = -2A\sigma^2\cos\pi - (-2A\sigma^2\cos 0) = 4A\sigma^2 \qquad (7\text{-}47)$$

式中，E_{ex} 为两原子自旋交换能；σ 是电子以 $\hbar$ 为单位的电子自旋角动量；A 为交换积分。但是如果在 n 个等距的原子面间逐步地均匀转向，如图 7-18a 所示，相邻两个原子之间的交换作用能 $E_{ex(\varphi)}$则可以写成

$$E_{ex(\varphi)} = -2A\sigma^2\cos\varphi \qquad (7\text{-}48)$$

式中，φ 为相邻两个原子磁矩间的夹角，$\varphi = \pi/n$。当 φ 很小时，$\cos\varphi \approx 1-\varphi^2/2$。这

时相邻两原子自旋的交换能的变化$(\Delta E_{ex})_\varphi$为

$$(\Delta E_{ex})_\varphi = E_{ex(\varphi)} - E_{ex(0)} = -2A\sigma^2\cos\varphi - 2A\sigma^2\cos 0 \approx A\sigma^2\left(\frac{\pi}{n}\right)^2 \quad (7\text{-}49)$$

在 $n+1$ 个自旋磁矩的转向中，交换能的总变化 ΔE_{ex} 为

$$\Delta E_{ex} = \sum_1^n (\Delta E_{ex})_\varphi = A\sigma^2\pi^2/n \quad (7\text{-}50)$$

比较式（7-47）和式（7-50），后一种情况比前一种情况的交换能低得多。因此，畴壁中的原子磁矩必然是逐步转向。畴壁是原子磁矩由一个磁畴的方向逐步转向相邻磁畴的方向的过渡区。在畴壁内的交换能、磁晶各向异性能和磁弹性能都可能比磁畴内高，所高出的部分的能量称畴壁能，用 E_ω 表示。畴壁单位面积的能量称畴壁能密度，用 γ_ω 表示。由式（7-50）可知，如果只考虑交换能，则在畴壁内相邻原子磁矩的方向改变越小，交换能越小，即交换能使畴壁无限加宽。但是事实上这是不可能的。因为 n 越大，就有更多的原子磁矩偏离易磁化方向，使磁晶各向异性能增加。磁晶各向异性能力图使畴壁变薄。综合考虑以上两个方面的因素，使总能量为最小，可以求出畴壁能密度 γ_ω 和畴壁厚度 δ 为

$$\delta = \pi\sqrt{A_1/K_1} \quad (7\text{-}51)$$

$$\gamma_\omega = 2\pi\sqrt{A_1K_1} \quad (7\text{-}52)$$

式中，$A_1 = A\sigma^2/a$，a 为点阵常数；K_1 为磁晶各向异性常数。考虑当材料内部存在内应力时，由于应力也要引起应力各向异性，根据式（7-46），可以将应力各向异性和磁晶各向异性同样考虑，求得总能量最小时的畴壁能密度 γ_ω 和畴壁厚度 δ 分别为

$$\delta = \pi\sqrt{A_1/[K_1 + (3/2)\lambda_s\sigma']} \quad (7\text{-}53)$$

$$\gamma_\omega = 2\pi\sqrt{A_1[K_1 + (3/2)\lambda_s\sigma']} \quad (7\text{-}54)$$

这里，λ_s 为饱和磁致伸缩系数；σ'为内应力。

可见畴壁厚度与材料的 K_1、A_1、λ_s、σ'等参量有关。K_1 越大，δ 越小，γ_ω 越大。在 Fe-Ni 合金中，K_1 很小，如果内应力也很小的话，则畴壁厚度可相当得大，这时畴壁内相邻原子间磁矩的角度 φ 仅有 0.18°～1.8°，磁矩的分布近似具有连续性，这种畴壁称为连续性的畴壁模型。在六方结构的 Co 和 $SmCo_5$ 等金属与合金中，由于 K_1 很大，δ 很小，γ_ω 很大。其 φ 角可达 6°～180°，并且 φ 角的分布是不均匀的。这种畴壁称为非连续畴壁模型。窄畴壁对材料磁特性有重要的影响。

7.6.2　磁畴

畴结构受到畴壁能 E_ω、磁晶各向异性能 E_K、磁弹性能 E_σ 和退磁场能 E_d 的制

约，其中退磁场能将是铁磁体分成畴的动力。其他能量将决定磁畴的形状、尺寸和取向。为简便起见，考虑一个边长为 1cm ×1cm×0.5cm 的方块形单晶铁，它是立方晶体，$K_1 > 0$。如果它是一个如图 7-19a 所示的单畴体，无畴壁，则 $E_\omega = 0$，且磁矩方向排列在铁的易极化方向[001]（同方向或反方向），那么 $E_K = 0$。若不考虑磁弹性能，即 $E_\sigma = 0$，方块形状决定的退磁场能 E_d 就是总能量 E^a，即

$$E^a = E_d = V\left(\frac{1}{2}NM_s{}^2\mu_0\right) \tag{7-55}$$

式中，V 是铁块体积。方块形状铁磁体的退磁因子接近球体的退磁因子，即 $N = 1/3$，铁的 $M_s = 1.73\times10^{-6}\text{A·m}^{-1}$，$V = 5\times10^{-7}\text{m}^3$，代入式（7-55），$E_d = 0.313$ J。

如果将它依次分成 2 个畴，4 个畴，……，n 个畴，如图 7-19b 所示，它的退磁场能可以近似地看作图 7-19a 的 $1/n$，这样分畴越多，退磁能就越低。但是从另一方面考虑，分畴越多，磁畴壁越多，畴壁能越高。当分为 n 个磁畴时，有 $n-1$ 块畴壁，这时的总能量 E^b（磁矩都在铁的易极化方向[001]，$E_K = 0$）为

$$E^b = V\left(\frac{1}{2}NM_s{}^2\mu_0\right)\Big/ n + \gamma_{\omega(180)}(n-1)S_{180} \tag{7-56}$$

式中，S_{180} 是 180°畴壁面积；$\gamma_{\omega(180)}$是 180°畴壁能密度。设 $n = 20$，则 $E^b = 0.016$ J，是 E^a 的 1/20。

如果形成图 7-19c 那样的封闭畴，它由 4 个三角畴和两块位于{011}面的 90°畴壁来组成。在畴壁内磁通是连续的，方块形铁磁体表面不会出现磁荷，则退磁场能 E_d 为零。[100]、[010]、[001]方向都是易磁化方向，所以全部磁畴中的磁矩的方向都排列在易磁化方向，$E_K = 0$，但这时 4 个三角畴都要沿自己的易磁化方向伸长，出现了由应力产生的磁弹性能。其总能量 E^c 为

$$E^c = E_\sigma + \gamma_{\omega(90)}S_{90} = \frac{V\lambda_{100}^2 C_{11}}{2} + \gamma_{\omega(90)}S_{90} \tag{7-57}$$

式中，S_{90} 是 90°畴壁的总面积，$S_{90} = 1.41\times10^{-4}\text{m}^2$；$\gamma_{\omega(90)}$是 90°畴壁能密度，$\gamma_{\omega(90)} = 1.07\times10^{-3}\ \text{J·m}^{-2}$；$\lambda_{100}$ 是沿<100>晶轴方向磁化到饱和时的饱和磁致伸缩系数，$\lambda_{100} = 2.07\times10^{-5}$；$C_{11} = 2.41\times10^{-12}\text{N·m}^{-2}$。则总能量 $E^c = 25.7\times10^{-5}$ J，可见出现封闭畴后，方块形铁的能量大大降低。

如果形成图 7-19d 那样的封闭畴，和图 7-19c 相比，三角畴的体积和 S_{90} 减少了，但是出现了 180°畴壁，其总能量 E^d 为

$$E^d = \frac{V_c\lambda_{100}^2 C_{11}}{2} + \gamma_{\omega(180)}S_{180} + \gamma_{\omega(90)}S_{90} \tag{7-58}$$

式中，V_c 是封闭畴体积；S_{180}、S_{90} 分别 180°和 90°畴壁的总面积。假定出现了 $n=8$ 块的片状畴，则畴宽 $\delta = 1.25\times10^{-3}\text{m}$，$V_c = 0.5\delta^2\times0.005n = 3.13\times10^{-8}\text{m}^3$，$S_{180} = (n-1)(l-2h)\ \times0.005 = 3.06\times10^{-4}\text{m}^2$，$S_{90} = 2n[\delta/\cos(\pi/4)] = 1.41\times10^{-4}\text{m}^2$，总能量为

$E^d = 16 \times 10^{-6}$ J。在图 7-19 的 4 种畴结构中，图 7-19d 是能量最低的。而实际方块形的立方单晶铁存在的畴结构是与图 7-19d 的畴结构相同，说明方块形单晶铁中封闭畴结构比图 7-19b 所示的片状畴结构的能量更低，但是具体形成的封闭畴的数量，要进一步计算才可知。

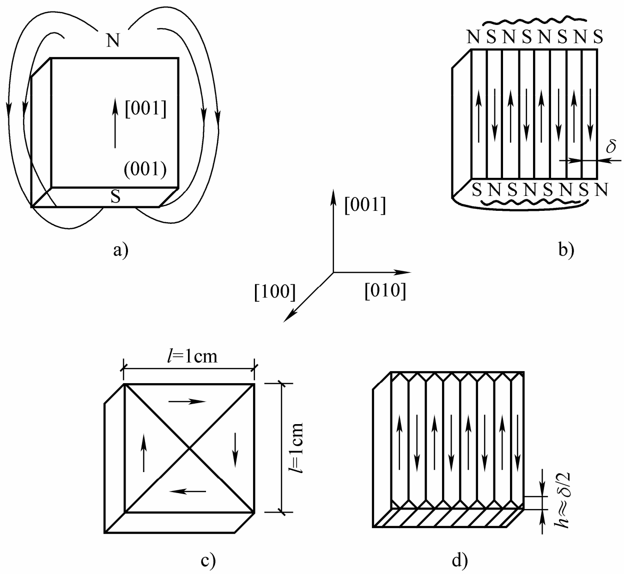


图 7-19　边长为 lcm×lcm×0.5cm 的方块形单晶铁的可能畴结构［正面是(001)面］

a）单畴体　b）多分畴　c）封闭畴　d）多畴的封闭畴

实际材料中的畴结构，还要受到材料的尺寸、晶界、应力、掺杂和缺陷等的影响，因此实际材料的畴结构是相当复杂的。例如为减少退磁场能，在掺杂物附近出现三角畴或钉状畴，如图 7-20a、b 所示。在多晶体中，畴壁一般不能穿过晶粒边界，如图 7-20c 所示。

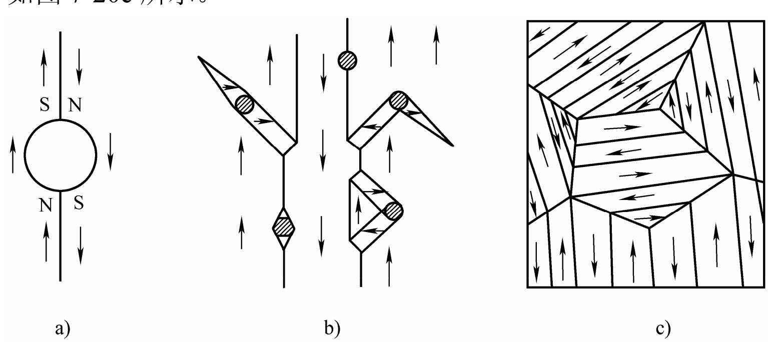


图 7-20　空洞、掺杂和晶界对磁畴结构的影响

a）空洞　b）掺杂　c）晶界

7.6.3 单畴结构

铁磁体的块度对磁畴结构也有很大影响。如果材料的线度非常小，以至材料形成单畴时的退磁能小于形成多畴时的畴壁能，磁性材料就以单畴存在。例如一个半径为 R 的立方单晶体结构的球（体积为 V）。设磁晶各向异性常数 $K_1>0$，且 λ_s 和应力 σ 可以忽略，于是如果分成图 7-21a 所示的 4 块封闭畴结构时，球体中其他能量都为 0，只有畴壁能，总能量是 $E^a=\gamma_{90}\times2\pi R^2$。如果不分畴，如图 7-21b 所示，球体中其他能量都为 0，只有退磁能，总能量是 $E^b = V\cdot(0.5NM_s^2\mu_0) = 8\mu_0\pi^2 M_s^2R^3/9$。当 $E^a = E^b$ 时，求出的球的半径称为单畴体的临界尺寸 R_c，单晶体球体的 $R>R_c$ 时，则分畴的情况下能量最低，以多畴体存在；当 $R<R_c$ 时，则不分畴的情况下能量最低，以单畴体存在，如图 7-21c 所示。磁记录用的磁粉和一些永磁材料都是单畴结构。

另外如果将单畴体的临界尺寸继续减小到一定程度后，由于表面与体积比大大增加，热振动能可能和微粒的磁晶各向异性能相当，这时微粒的磁矩不能固定地沿着易磁化方向排列，它随热振动自由改变，单畴微粒体就转变为超顺磁体。由单畴体转变为超顺磁体的临界尺寸 Dp 称为超顺磁体的临界尺寸。磁性流体中使用的磁性微粒一般具有超顺磁性。

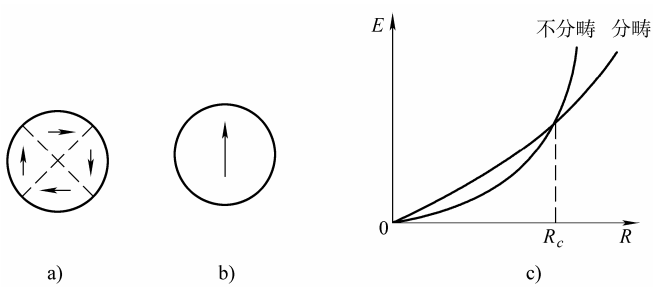


图 7-21 立方单晶铁磁体球状颗粒（$K_1>0$）的磁畴结构以及能量变化

a）分畴 b）不分畴 c）能量变化

第 8 章　第一性原理计算

第一性原理计算（First−principles calculation）就是从电子结构出发，应用量子力学理论，只借助基本常量和某些合理的近似进行的计算。这种计算如实地把固体作为电子和原子核组成的多粒子系统，求出系统的总能量。根据总能量与电子结构和原子核构型的关系，确定系统的状态，预测材料的性能。本章简要介绍了第一性原理的理论基础和计算方法、常用的第一性原理计算软件以及在材料科学中的研究应用等。

8.1　引言

第一性原理是指在绝热近似和单电子近似的基础上，在计算中仅仅使用普朗克常量、电子质量和电量三个基本物理常数，以及原子的核外电子排布，而不借助任何可调节的经验参数，通过自洽计算来求解薛定谔方程。如果只考虑一个电子，而把其他电子对它的作用近似地处理成某种平均势场，这样就可以把多电子问题简化成单电子问题，这种近似称为单电子近似，也称为平均场近似。哈特里−福克（Hartree−Fork）近似是平均场近似的一种，它是把所要讨论的电子视为在离子势场和其他电子的平均势场中的运动。但是，哈特里−福克近似的程度过大，因为它忽略了电子之间的交换和相关效应，使得计算的精度受到了一定的限制。为了解决这一问题，P. Hohenberg 和 W. Kohn 于 1964 年提出了密度泛函理论（density functional theory，DFT），这一理论巧妙地将电子之间的交换相关势表示为密度泛函，使得薛定谔方程在考虑了电子之间的复杂相互作用后，依然可以利用自洽的方法求解。DFT 认为粒子的哈密顿量取决于电子密度的局域值，由此可以得出局域密度近似方法（Local Density Approximation，LDA）。LDA 将多电子系统转化为单电子，用此理论对各类半导体和金属材料和结合能、晶格常数、体变模量的计算结果与实验符合得很好，成为近年来电子理论中的一项重要的成就。由于固体实验提供了大量的实验数据，使得能带理论的实用性得到验证，用此理论也由对固体定性的普遍性规律探讨到对具体材料复杂能带结构的计算。

有多种手段可以实现第一性原理的计算，目前广泛应用的是赝势（Pseudo−potential）方法。赝势方法的核心思想是将内层电子对孤立正电荷库仑势的屏蔽作用，等价为一个在近原子核区域平缓变化的有效势。引入赝势后，价电子所受作

用力被有效势代替。这样，在实际计算过程中所用的赝势和对应的赝波函数只要极少的平面波就足以展开。

第一性原理计算，因其不含任何可调（经验或拟合）参数，就可以通过求解薛定谔（Schrödinger）方程，得到材料的电子结构，预测材料的组分、结构与性能，设计具有特定性能的新材料，从而在计算材料学中占有其他计算方法所无法取代的地位。第一性原理计算，一方面对我们的真实实验进行有效的补充，通过计算使被模拟体系的特征和性质在理论上更加全面地显示出来，从而更全面了解事物；另一方面，与真实的实验相比，第一性原理计算也能指导实验，使我们更快更合理地设计出符合要求的实验。因此，第一原理方法是一种真正意义上的预测方法。第一原理方法的基本计算结果为体系总能量以及电荷分布（电荷密度，态密度），很多更加实用的量如弹性常数，点及面缺陷的形成能均可从这些量推演而来。更进一步，借助于某些统计力学，第一原理计算还可为相变及合金中的相图本质提供有益的启示。

第一性原理方法的不足之处有以下几点：

1）所处理的体系较简单，一般在 1000 个原子以内（大多数低于 100 个原子）。

2）计算所用时间长。

3）体系的许多性能得不到计算，常要外推或用基态的性质代替。

8.2 第一性原理的基础理论

第一原理计算有着较完善的理论基础，在求解体系的薛定谔方程的过程中不涉及任何经验参数，所要输入的只是原子的核电荷数和一些模拟的环境参量。计算所求得的结果是体系薛定谔方程的本征值和本征函数（波函数），有了这两项结果从理论上讲就可以推导出体系的所有特性。下面介绍第一性原理计算的理论基础。

8.2.1 多原子体系的薛定谔方程（Schrödinger function）

对于固体材料系统，必然是个极大的多电子体系，要确定固体电子的能级，其出发点就是处理组成固体的多原子体系的薛定谔方程

$$\hat{H}\psi(\boldsymbol{r},\boldsymbol{R}) = E\psi(\boldsymbol{r},\boldsymbol{R}) \tag{8-1}$$

其中，$\boldsymbol{r}$ 表示所有电子坐标$\{\boldsymbol{r_i}\}$的集合，$\boldsymbol{R}$ 表示所有原子核坐标$\{\boldsymbol{R_i}\}$的集合。在不考虑其他外场时，哈密顿量 $\hat{H}$ 包括原子核的动能 T_N 和电子的动能 T_e、电子间相互作用势 V_e、原子核间相互作用势 V_N 以及电子−原子相互作用势 V_{N-N}。

$$\hat{H} = T_e(\boldsymbol{r}) + V_e(\boldsymbol{r}) + T_N(\boldsymbol{R}) + V_N(\boldsymbol{R}) + V_{e-N}(\boldsymbol{r},\boldsymbol{R}) \tag{8-2}$$

1. 玻恩–奥本海默（Born-Oppenheimer）近似

固体中的电子结构计算实际上是一个同时处理电子与原子核运动的问题。由于电子的质量远小于原子核的质量，故它的运动远快于核的热运动，电子处于高速运动中，而原子核只是在它们的平衡位置附近振动。电子能绝热于原子的运动，而原子核只能缓慢地跟上电子分布的变化。基于这一特性，M. Born 与 J. E. Oppenheimer 建议可将电子与原子核的运动分开处理，称为玻恩–奥本海默（Born–Oppenheimer）近似（或绝热近似）。其主要思想是：

1）将固体整体的平移、转动（外运动）和核的振动运动分离出去。

2）考虑电子运动时，将坐标系原点设定在固体质心上，并令其随固体整体一起平移和转动，同时令各原子核固定在它们振动运动的某一瞬时位置上。

3）考虑核的运动时则不考虑电子在空间的具体分布。

因此，在绝热近似条件下，电子和原子核的运动可以分开，固体中多粒子系统的薛定谔方程的解可以写成原子核运动波函数 $\phi(\boldsymbol{R})$ 和电子运动波函数 $\Phi(\boldsymbol{r}, \boldsymbol{R})$之乘积。其中多电子波函数 $\Phi(\boldsymbol{r}, \boldsymbol{R})$由原子核的瞬时坐标 $\boldsymbol{R}$ 下的多电子哈密顿量 $\hat{H}_0$ 决定。用外势场 V_{ext} 表示多电子体系所受到的外部作用，$\hat{H}_0$ 包括电子–原子核作用和其他来自体系外部的势场，则可得到

$$\hat{H}_0\phi_n(\boldsymbol{r},\boldsymbol{R}) = E_n\phi_n \tag{8-3}$$

其中，$\hat{H}_0 = T_e(\boldsymbol{r}) + V_e(\boldsymbol{r}) + V_{ext}$。

用 $V(\boldsymbol{r}, \boldsymbol{R})$代表薛定谔方程中的势能项，可得

$$V(\boldsymbol{r},\boldsymbol{R}) = \sum_{p<q}\frac{Z_pZ_q}{\boldsymbol{R}_{pq}} + \sum_{i<k}\frac{1}{\boldsymbol{r}_{ik}} - \sum_{pj}\frac{Z_p}{\boldsymbol{r}_{pj}} \tag{8-4}$$

则分离变量后得到的电子运动方程为

$$-\frac{1}{2}\sum\nabla_i^2\Phi(\boldsymbol{r},\boldsymbol{R}) + V(\boldsymbol{r},\boldsymbol{R})\Phi(\boldsymbol{r},\boldsymbol{R}) = E(\boldsymbol{R})\Phi(\boldsymbol{r},\boldsymbol{R}) \tag{8-5}$$

原子核的运动方程为

$$-\frac{1}{2}\sum\nabla_p^2\phi(\boldsymbol{r},\boldsymbol{R}) + E(\boldsymbol{R})\phi(\boldsymbol{R}) = E\phi(\boldsymbol{R}) \tag{8-6}$$

方程（8-5）为在某种固定的核的位置时体系的电子运动方程，方程（8-6）为原子核的运动方程。$E(\boldsymbol{R})$在方程（8-5）中代表原子核固定时电子的总能量，而在方程（8-6）中作为核运动的势能出现，是一个常数。因此，在计算势能曲线或曲面时，$E(\boldsymbol{R})$又称为势能。E 为体系总能量，包括电子总能量和原子核的振动能量。这两个方程表明，原子核在势函数为 $E(\boldsymbol{R})$的势阱中运动。解电子系统的薛定谔方程（8-5），得到电子波函数 $\Phi(\boldsymbol{r}, \boldsymbol{R})$和电子总能量 $E(\boldsymbol{R})$。以 $E(\boldsymbol{R})$作为核运动的势函数，解核的薛定谔方程（8-6），得到核波函数 $\phi(\boldsymbol{R})$。对于分子，便可得到分子的平动、转动和振动态。

Born−Oppenheimer 近似又称为“绝热近似”，它认为电子总是不断地适应核位置的变化，或者说跟随着核且始终围绕着核来运动，运动过程中不会发生从一个势能面至另一个势能面的跃迁，即吸收或放出能量发生电子能级跃迁。

由于涉及电子间的多体相互作用问题以及电子−原子核之间的相互作用，而固体往往又具有非常大的体积，所含电子数目庞大，因此这样复杂的方程难以严格求解，需要作进一步的近似。

2. 哈特里−福克（Hartree-Fock）近似

通过绝热近似后，把电子和原子核的运动分开，得到了多电子体系的薛定谔方程。问题的关键在于多个电子的严格波函数难于表达和计算。Kohn 甚至断言，当体系电子数超过 1000 个以后，多电子波函数 $\Phi(x_1, x_2, \cdots, x_N)$ 是一个不合理的科学概念，所谓“不合理”，其意义是指波函数无法用当今的计算机和存储技术准确的计算以及正确的记录下来。而 $1m^3$ 的实际固体中却约有 10^{29} 个电子！

对于多电子体系，上述简化后的定态薛定谔方程仍然不可能严格求解，原因是多电子势函数中包含了 $\boldsymbol{r}_{ik}^{-1}$ 形式的电子间排斥作用算符，不能分离变量。为了解决这个问题，人们对电子相互作用项作了种种近似。在密度泛函理论形成之前，Hartree 和 Fock 先后分别提出了求解方程的自洽场近似方法，简称哈特里−福克近似。

Hartree 提出：如果完全忽略电子−电子相互作用，每个电子的状态可分别用一个单电子波函数 $\Phi_i(x_i)$来表示。N 个电子体系的总波函数写成 N 个单电子波函数的乘积

$$\Phi(x_1, x_2, \cdots, x_N) = \Phi_1(x_1)\Phi_2(x_2), \cdots, \Phi_N(x_N) \tag{8-7}$$

这种近似称为哈特里（Hartree）近似。其中每一个单电子函数 $\Phi_i(x_i)$只与一个电子的坐标 x_i 有关，将其代入电子的多体薛定谔方程，利用多电子系统总能量对试探单电子波函数的变分极小，得到著名的哈特里（Hartree）方程

$$[-\nabla^2 + V(\boldsymbol{r}) + \sum_{i' \neq i} \int \frac{|\Phi_{i'}(\boldsymbol{r}')|^2}{|\boldsymbol{r}' - \boldsymbol{r}|} \mathrm{d}\boldsymbol{r}'] \Phi_i(\boldsymbol{r}) = E_i \Phi_i(\boldsymbol{r}) \tag{8-8}$$

其中，V 是体系原子核产生的势场，$\sum_{i' \neq i} \int |\Phi_{i'}(\boldsymbol{r}')|^2 \big/ |\boldsymbol{r}' - \boldsymbol{r}| \mathrm{d}\boldsymbol{r}'$ 算符项又称为哈特里（Hartree）项，这是一个势场算符，代表其他电子（$i' \neq i$）共同产生的一个“合”平均场。方程（8-8）描述了 $\boldsymbol{r}$ 处单电子在晶格势 V 和其他电子的平均势中的运动，即将电子间的库仑排斥平均化，每个电子均视为在核库仑场与其他电子对核电子的平均势相叠加而成的势场中运动，从而单个电子的运动特性只取决于其他电子的平均密度分布（即电子云），而与后者的瞬时位置无关。

这个近似隐含的物理模型是一种“独立电子模型”，有时又称为“单电子近似”。虽然哈特里多电子波函数中每个量子态不同，满足不相容原理，但并没有考虑到

电子波函数的交换反对称性。Fock 注意到这一点，用单电子波函数的斯莱特（Slater）行列式代替多体波函数

$$\Phi(x_1,x_2,\cdots,x_N)=\frac{1}{\sqrt{N!}}\begin{vmatrix}\Phi_1(x_1) & \Phi_2(x_1) & \cdots & \Phi_N(x_1)\\ \Phi_1(x_2) & \Phi_2(x_2) & \cdots & \Phi_N(x_2)\\ \vdots & \vdots & & \vdots\\ \Phi_1(x_N) & \Phi_2(x_N) & \cdots & \Phi_N(x_N)\end{vmatrix} \tag{8-9}$$

根据数学完备集理论，体系状态波函数 ψ 应该是无限个斯莱特行列式波函数的线性组合，即把式（8-9）中的单个行列式波函数记为 D_p，则通过变分处理能得到玻恩−奥本海默近似下的任意精确的能级和波函数。这种近似称为福克（Fock）近似。福克近似所采用的波函数既满足泡利不相容原理又满足交换反对称性：两电子具有相同状态相当于斯莱特行列之中的两列相等，则波函数 ψ=0；交换两个电子相当于交换斯莱特行列式的两行，使波函数 ψ 改变符号。福克近似的实质是用归一化的单电子波函数的乘积线性组合成具有交换反对称性的函数作为多电子系统的波函数。

采用哈特里近似和福克近似，根据变分原理，由最佳单电子波函数 Φ_i 构成的体系状态波函数 ψ 一定给出系统能量 E 的极小值。将 E 对 Φ_i 作变分，以 E_i 为拉格朗日乘子，得到单电子波函数应满足的微分方程

$$[-\nabla^2+V(\boldsymbol{r})]\Phi_i(\boldsymbol{r})+\sum_{i'\neq i}\int \mathrm{d}\boldsymbol{r}'\frac{|\Phi_{i'}(\boldsymbol{r}')|^2}{|\boldsymbol{r}'-\boldsymbol{r}|}\Phi_i(\boldsymbol{r})+\sum_{i'\neq i,\|}\int \mathrm{d}\boldsymbol{r}'\frac{\Phi_{i'}^*(\boldsymbol{r}')\Phi_i(\boldsymbol{r}')}{|\boldsymbol{r}'-\boldsymbol{r}|}\Phi_{i'}(\boldsymbol{r})=E_i\Phi_i(\boldsymbol{r}) \tag{8-10}$$

方程（8-10）表示的单电子方程称为哈特里−福克（Hartree−Fock）方程。哈特里−福克方程左边第一项是单电子动能和原子核对单电子的作用项，第二项是电子库仑相互作用项，第三项是电子交换相互作用项，“$\|$”表示求和只对与 Φ_i 有平行自旋的 Φ_i 进行，即只包含自旋平行的电子间的交换作用。

哈特里−福克方程只能用迭代法求解。先设哈特里−福克方程的初解为一组单电子态$\{\Phi_i\}$，根据所设单电子态求出势函数，解方程得到更好的解$\{\Phi_i\}$，重复这一过程，直到$\{\Phi_i\}$在所考虑的计算精度内不再变化，即由单电子态决定的势场与由势场决定的单电子态之间达到自洽，这就是哈特里−福克自洽场近似方法（简称 HF 方法）。利用哈特里−福克近似，我们将多电子薛定谔方程简化为单电子有效势方程。由此可见，哈特里−福克近似的实质是将每个电子的运动近似成单个电子在一个有效势场中的独立运动，这种近似称为单电子近似。在哈特里−福克近似中，有效势包含了原子核对电子的静电吸引作用、电子与电子的库仑排斥作用和电子与电子的交换相互作用。

这个方法最大的优点就是它计算结果的精确性，它是严格意义上的从头（ab−initio）方法。但也存在现在还难以克服的困难，就是此计算方法的计算量随着电

子数的增多呈指数增加。因此，这种计算对计算机的内存大小和 CPU 的运算速度有非常苛刻的要求，它使得对具有较多电子数的计算成为不可能，如含有过渡元素或重金属元素体系的计算。一般此方法多用于轻元素的计算，如 C、H、O 以及 N 等，多用于化学计算，这在很大程度上也是导致密度泛函理论产生的原因。

8.2.2 密度泛函理论（DFT）

20 世纪 60 年代，Hohenberg、Kohn 和 Sham（沈吕九）提出了密度泛函理论，简称 DFT。密度泛函理论将多电子波函数和薛定谔方程用简单的电荷密度 $\rho(\boldsymbol{r})$和相应的计算方案来代替，不但给出了将多电子体系简化为单电子问题的理论基础，同时也成为当今分子和固体电子结构和总能量计算的有力工具。因此，密度泛函理论是多粒子系统理论基态研究的重要方法。DFT 同分子动力学方法相结合，在材料设计、合成、模拟计算和评价等许多方面有明显的进展，成为计算材料科学的重要基础和核心技术。近年来，用 DFT 的工作以指数规律增加，现在已经大大超过用 HF 方法研究的工作。DFT 适应于大量不同类型的应用，因为电子基态能量与原子核位置之间的关系可以用来确定分子或晶体的结构，而当原子不在它的平衡位置时，DFT 可以给出作用在原子核位置上的力。因此，DFT 可以解决原子分子物理中的许多问题，如电离势的计算、振动谱研究、化学反应问题、生物分子的结构和催化活性位置的特性、等等。在凝聚态物理中，如材料电子结构和几何结构、固体和液态金属中的相变等。DFT 的另一个优点是，它提供了第一性原理或从头算的计算框架，在这个框架下可以发展各式各样的能带计算方法。下面对密度泛函理论进行介绍。

1．霍恩伯格–科恩（Hohenberg-Kohn）定理

1927 年，H. Thomas 和 Fermi 各自提出以电子密度的函数来表示能量的理论，这样当然大大减少了计算中的变量，但是由于缺乏对动能项的精确处理，这个早期的理论有点粗糙，主要侧重于化学方面，甚至有时会给出分子不会形成键结的计算结果，这个结论阻碍了托马斯–费米（Thomas–Fermi）理论的进一步发展。到了 20 世纪 60 年代，从电子密度出发的概念给出了 Kohn 新的想法，1964 年 Hohenberg 和 Kohn 发表了后来极著名的霍恩伯格–科恩（Hohenberg–Kohn，HK）定理，对于无简并的基态：

定理一：不计自旋的全同费米子系统的基态能量是粒子数密度 $\rho(\boldsymbol{r})$函数的唯一泛函（除了一附加常数外）。

定理二：能量泛函 $E[\rho]$在粒子数保持不变的条件下对正的粒子数密度函数取极小值，并等于基态能量 E_0，即在条件 $\rho(\boldsymbol{r})>0$ 和 $\int\rho(\boldsymbol{r})\mathrm{d}\boldsymbol{r}=N$ 下，满足 $\delta E[\rho]=0$ 时的 ρ 为基态密度泛函，此时的能量等于基态能量 E_0。

这两个定理统称为霍恩伯格−科恩定理。定理一说明，粒子数密度函数是确定多粒子系统基态物理性质的基本变量，多粒子系统的所有基态物理性质，如能量、波函数以及所有算符的期待值，都由粒子数密度函数唯一确定。定理二表明，如果得到了基态粒子数密度函数，就能确定能量泛函的极小值，并且这个极小值等于基态的能量，因此，能量泛函对粒子数密度的变分是确定系统基态的途径。

对于给定的局域势 $v(\boldsymbol{r})$（表示外场的作用），多电子系统的能量是电子数密度（也称为电荷密度）$\rho(\boldsymbol{r})$的泛函，则对应的多电子系统的能量泛函 $E[\rho]$为

$$E[\rho]=F[\rho]+E_{ext}[\rho]+E_{N-N} \tag{8-11}$$

式中，$F[\rho]$是与外场无关的泛函，它包括电子的动能和电子之间的相互作用能

$$F[\rho]=T[\rho]+E_{e-e}[\rho]=T[\rho]+\frac{1}{2}\iint \mathrm{d}\boldsymbol{r}\mathrm{d}\boldsymbol{r}'\frac{\rho(\boldsymbol{r})\rho(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}+E_{xc}[\rho] \tag{8-12}$$

$E_{ext}[\rho]$是局域势 $v(\boldsymbol{r})$所表示的外场对电子的作用能

$$E_{ext}[\rho]=\int \mathrm{d}\boldsymbol{r}v(\boldsymbol{r})\rho(\boldsymbol{r}) \tag{8-13}$$

E_{N-N} 是原子核之间的排斥能

$$E_{ext}[\rho]=\sum_{i<j}\frac{Z_iZ_j}{|\boldsymbol{R}_i-\boldsymbol{R}_j|} \tag{8-14}$$

为了说明 $F[\rho]$，在式（8-12）中将 $F[\rho]$分为三项，第一项和第二项是与无相互作用粒子模型对应的动能和库仑排斥能，第三项 $E_{xc}[\rho]$代表了所有未包含在无相互作用粒子模型中的相互作用能。这里所谓无相互作用是指一个电子的存在对其他电子没有影响，而实际上一个电子的存在对其他电子是有影响的。如果在 $\boldsymbol{r}$ 处存在一个电子，那么在 $\boldsymbol{r}'$处的电子数密度将不再是 $\boldsymbol{r}$ 处无电子时的 $\rho(\boldsymbol{r})$，这表明电子间除了库仑排斥作用还存在其他的相互作用，这种相互作用包括自旋平行电子间的交换相互作用和自旋反平行电子间的关联相互作用。因此，$E_{xc}[\rho]$称为交换关联相互作用能（简称交换关联能），它包含了电子间相互作用的全部复杂性。$E_{xc}[\rho]$也是电子数密度 ρ 的泛函，可分成交换能 $E_x[\rho]$和关联能 $E_c[\rho]$两部分，即 $E_{xc}[\rho]=E_x[\rho]+E_c[\rho]$。

根据霍恩伯格−科恩定理，如果能得到能量泛函 $E[\rho]$，将 $E[\rho]$对电子数密度 ρ 变分，就可以确定系统的基态和所有的基态性质。但要得到能量泛函 $E[\rho]$，必须解决以下三个问题：

1）粒子数密度函数 $\rho(\boldsymbol{r})$ 的确定。

2）动能泛函 $T[\rho]$的确定。

3）交换关联泛函 $E_{xc}[\rho]$ 的确定。

W. Kohn 和 L.J. Sham（沈吕九）提出了解决第一和第二个问题的方法，并由此得到了 Kohn−Sham 方程。对于第三个问题，一般采用局域密度近似来解决。

2. 科恩–沈［吕九］（Kohn-Sham）方程

上述霍恩伯格–科恩定理说明了电荷密度是确定多电子体系基态物理性质的基本变量，能量泛函对电荷密度函数的变分是确定系统的途径，但由于多电子体系中电子间的相互作用，严格确定相互作用系统的密度泛函 $E_{xc}[\rho]$非常困难。

由于对有相互作用电子系统的动能项一无所知，W. Kohn 和 L.J. Sham 提出：

1）假定动能泛函 $T[\rho]$可用一个已知的无相互作用电子系统的动能泛函 $T_S[\rho]$来代替，这个无相互作用电子系统与有相互作用电子系统具有相同的密度函数。

2）用 N 个单电子波函数 $\Phi_i(\boldsymbol{r})$构成密度函数 $\rho(\boldsymbol{r})$

$$\rho(\boldsymbol{r}) = \sum_{i=1}^{N} \left|\Phi_i(\boldsymbol{r})\right|^2 \tag{8-15}$$

根据霍恩伯格–科恩定理一，如果 $\rho(\boldsymbol{r})$确定，$V_S(\boldsymbol{r})$就唯一确定，因此无相互作用多粒子体系的能量也是 $\rho(\boldsymbol{r})$的单值函数

$$E_S[\rho(\boldsymbol{r})] = \sum_{i}^{N} \varepsilon_i = T_S[\rho(\boldsymbol{r})] + \int V_S(\boldsymbol{r})\rho(\boldsymbol{r})\mathrm{d}\boldsymbol{r} \tag{8-16}$$

其中，$T_S[\rho(\boldsymbol{r})]$是无相互作用多粒子体系的动能泛函

$$T_S[\rho(\boldsymbol{r})] = \sum_{i=1}^{N} \int \mathrm{d}\boldsymbol{r}\Phi_i^*(\boldsymbol{r})[-\nabla^2]\Phi_i(\boldsymbol{r}) \tag{8-17}$$

将能量泛函 $E[\rho]$对 ρ 的变分用对 $\Phi_i(\boldsymbol{r})$的变分代替，以 ε_i 为拉格朗日乘子，变分后得到

$$\left(\frac{1}{2}\nabla^2 + V_{KS}(\boldsymbol{r})\right)\Phi_i(\boldsymbol{r}) = \varepsilon_i\Phi_i(\boldsymbol{r}) \tag{8-18}$$

其中，

$$\begin{aligned} V_{KS}[\rho(\boldsymbol{r})] &= v(\boldsymbol{r}) + V_{coul}[\rho(\boldsymbol{r})] + V_{xc}[\rho(\boldsymbol{r})] \\ &= v(\boldsymbol{r}) + \int \mathrm{d}\boldsymbol{r}' \frac{\rho(\boldsymbol{r}')}{\left|\boldsymbol{r} - \boldsymbol{r}'\right|} + \frac{\delta E_{xc}[\rho(\boldsymbol{r})]}{\delta\rho(\boldsymbol{r})} \end{aligned} \tag{8-19}$$

这样，对于单粒子波函数 $\Phi_i(\boldsymbol{r})$，也得到了与哈特里–福克方程相似的单电子方程，它包括局域势 $v(\boldsymbol{r})$、库仑排斥势 $V_{coul}[\rho(\boldsymbol{r})]$和交换关联势 $V_{xc}[\rho(\boldsymbol{r})]$。式（8-15）、式（8-18）和式（8-19）一起成为科恩–沈［吕九］方程。现在，基态的电子数密度函数可由解方程（8-18）得到的 $\Phi_i(\boldsymbol{r})$按式（8-15）构成。根据霍恩伯格–科恩定理，这样得到的电子数密度可精确地确定该系统基态的能量、波函数及各物理量的期待值

科恩–沈［吕九］方程的核心是，用无相互作用电子系统的动能代替有相互作用粒子系统的动能，而将有相互作用电子系统的全部复杂性归入交换关联相互作用泛函 $E_{xc}[\rho(\boldsymbol{r})]$中，从而导出形如式（8-18）的单电子方程。与哈特里–福克方程相比，密度泛函理论导出的单电子科恩–沈［吕九］方程是严格的，因为哈特里–福克方程只包含电子的交换相互作用，不包含电子的关联相互作用，而科恩–沈［吕九］方

程不仅包含电子的交换相互作用，而且包含电子的关联相互作用。

在密度泛函理论中，利用科恩–沈［吕九］方程可将多电子系统的基态特性问题转化成等效的单电子问题，这种计算方法与哈特里–福克自洽场近似法相似，但其结果更精确。然而，遗憾的是归纳了所有复杂相互作用的交换关联能泛函 $E_{xc}[\rho(\boldsymbol{r})]$ 是未知的，只有得到准确的、便于表达的 $E_{xc}[\rho(\boldsymbol{r})]$，才能求解科恩–沈［吕九］方程。科恩–沈［吕九］方程主要有以下几个特点。

1）通过引入 N 个单电子波函数，严格计算出了动能的主要部分，代价是需要求解 N 个方程。

2）除了更一般的局域势外，科恩–沈［吕九］方程与哈特里方程具有相似的形式，求解两种方程的计算量也相差不大，但比求解具有非定域势的 HF 方程要简单。

3）尽管哈特里（Hartree）、哈特里–福克（Hartree-Fock）和科恩–沈［吕九］方程都提供了一个多电子体系的单电子方法，但三者具有本质的差别，前两者一开始就引入了近似，而科恩–沈［吕九］原则是严格的，它可由变分原理严格推导。

至此，利用密度泛函理论计算多电子系统总能和电荷密度空间分布的方案可以付诸实施了。

8.2.3　交换关联能泛函

在整个密度泛函的理论中，由于交换关联能泛函 $E_{xc}[\rho(\boldsymbol{r})]$是未知的，具体计算中需要作适当的近似。显然密度泛函计算结果的精度，取决于交换关联势质量的好坏。

1. 局域密度近似（Local Density Approximation，LDA）

局域密度近似是实用中最简单有效的近似。它最早由 Slater 在 1951 年提出并应用，甚至早于密度泛函理论。这种近似假定某处的交换关联能只与该处的密度有关，且等于同密度的均匀电子气的交换关联能，因此可以把交换关联能近似写为

$$E_{xc}^{LDA}[\rho(\boldsymbol{r})] = \int \varepsilon_{xc}[\rho(\boldsymbol{r})]\rho(\boldsymbol{r})\mathrm{d}\boldsymbol{r} \tag{8-20}$$

其中，$\varepsilon_{xc}[\rho(\boldsymbol{r})] = \varepsilon_{xc}^{ham}[\rho(\boldsymbol{r})]$，是每个电子在密度为 $\rho(\boldsymbol{r})$的均匀电子气中的交换能和关联能之和。则交换关联势为

$$\begin{aligned} V_{xc}(\boldsymbol{r}) &= \frac{\delta E_{xc}[\rho(\boldsymbol{r})]}{\delta[\rho(\boldsymbol{r})]} = \frac{\delta[\rho(\boldsymbol{r})]\varepsilon_{xc}[\rho(\boldsymbol{r})]}{\delta[\rho(\boldsymbol{r})]} \\ &= \varepsilon_{xc}^{ham}[\rho(\boldsymbol{r})] + \rho(\boldsymbol{r})\frac{\mathrm{d}\varepsilon_{xc}^{ham}[\rho(\boldsymbol{r})]}{\mathrm{d}\rho(\boldsymbol{r})} \end{aligned} \tag{8-21}$$

如果考虑自旋不同的电子的密度 $\rho(\boldsymbol{r})$ 和轨道也不相同，相应的 LDA 就称为 LSDA 或者 LSD。

LDA 近似在大多数的材料计算中展示了巨大的成功。经验显示，LDA 计算原子游离能、分子解理能误差在 10%～20%；对分子键长、晶体结构可到 1%左右。

但是，局域密度泛函近似认为交换关联能函数完全是局域的，原则上忽略了某一点 $\boldsymbol{r}$ 附近电荷密度的非均匀性所带来的影响。虽然对于均匀电子气的交换关联能的计算结果较准确，但是对于与均匀电子气或者空间缓慢变化的电子气相差太远的系统，也就是在一些能量梯度很高的情况下，如方向性的成键，LDA 就难以得到精确的结果，往往过高估计键能。此外，分子的离解能和固体中的内聚能也常常被高估，虽然这种误差通常可在接受的精度范围内。因此对 LDA 需加以改进，使其适应性和精度提高。

2. 广义梯度近似（GGA）

迄今为止，已有数种尝试改进 LDA 的方案。其中较为成功的是广义梯度近似（GGA）。在 GGA 近似下，交换相关能是电子（极化）密度及其梯度的泛函。构造 GGA 交换相关泛函的方法分为两个流派。以 Becke 为首的流派认为“一切都是合法的”，所以人们可以以任何原因选择任何可能的泛函形式，而这种形式的好坏由实际计算来决定。通常，这样的泛函的参数由拟合大量的计算数据得到；以 Perdew 为首的流派主张尽量少引入拟合参数，通过渐进行为等物理限制来得到较好的泛函形式。基于这种理念构造的一个著名的 GGA 泛函是 PBE 泛函，也是现在用得最广泛的 GGA 泛函之一。除了密度梯度以外，在泛函中包含其他自变量（如动能密度）而得到的泛函称为 meta-GGA。泛函不同的 LDA 方案之间大同小异，但不同的 GGA 方案可能给出完全不同的结果。

GGA 的交换-关联能表达式与 LDA 类似，只是在其 $\varepsilon_{xc}[\rho(\boldsymbol{r})]$中加入了电子密度梯度，如

$$E_{xc}^{\mathrm{GGA}}[\rho(\boldsymbol{r})] = \int \mathrm{d}\boldsymbol{r}\, \rho(\boldsymbol{r}) \varepsilon_{xc}[\rho(\boldsymbol{r}), |\nabla \rho(\boldsymbol{r})|] \tag{8-22}$$

由于加入了一个非局域梯度项，GGA 从原理上来说应该更适合于处理非均匀体系。事实上，与 LDA 比较，GGA 只改进了原子交换关联能的计算，对价电子的电离能仅有小的变化。分子中的键长和固体中的晶格常数稍有增加，离解能和内聚能明显下降。然而至少对于较轻的元素，GGA 的结果一般与实验符合得很好，不仅是其价键和金属键，氢键和范德华键的键能计算都得到了改善。

应该指出的是，GGA 并非总是优于 LDA。例如，对于 4d 过渡金属的晶格常数和内聚能，GGA 的结果常常偏大，反不如 LDA 给出的精确。因此，很难说 GGA 与 LDA 谁更优越，只能由实际体系和计算的物理量来决定。由于 GGA 的计算量要高于 LDA，迄今为止 LDA 仍是使用最为广泛的交换关联能近似。

需要注意的是，如果交换关联能泛函显式包含轨道时，我们无法直接计算泛函微分，通常需要一种非直接的方法来求泛函极值。

3. 杂化密度泛函

哈特里–福克（Hartree−Fock）理论包含精确的交换关联能。因此，在 DFT 混合部分哈特里–福克交换可能有助于提高精度。通常，从科恩–沈［吕九］轨道得到的哈特里–福克交换被另外称为精确交换。最简单的杂化泛函可以写成

$$E_{xc} = aE_x^{\text{exact}} + (1-a)E_x^{\text{GGA}} + E_c^{\text{GGA}} \tag{8-23}$$

一个比较实用的杂化泛函是 B3LYP

$$E_{xc}^{\text{B3LYP}} = a_0 E_x^{\text{exact}} + (1-a_0)E_x^{\text{slater}} + a_c E_c^{\text{VMN}} + (1-a_c)E_c^{\text{LYP}} \tag{8-24}$$

杂化密度泛函在分子体系中被广泛应用，但较少应用到固体体系，因为固态计算中流行的平面波基组方法实现精确交换的效率太低。

8.2.4　赝势（Pseudopotential）

在求解固体的单电子波动方程时，用假想的势能代替离子实内部的真实势能，若不改变电子的能量本征值及其在离子实之间区域的波函数，则这个假想的势能叫赝势。利用赝势求出的价电子波函数叫赝波函数，赝波函数所满足的波动方程

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + V^{ps}\right]\psi_V^{ps} = E_V \psi_V^{ps} \tag{8-25}$$

式（8-25）又称为赝势方程。方程中 V^{ps} 是赝势，ψ_V^{ps} 是价电子的赝波函数，E_V 是价电子的能量。

在实际求解科恩–沈［吕九］方程时，由于原子核产生的势场项在原子中心处是发散的，波函数变化剧烈，需要用大量的平面波展开，因而计算变得十分困难。通常的做法是把空间区域按糕模（Muffin-tin）球划分，波函数在球内的部分由球面波展开，同时设定球外的原子势场为常数零。也可以用赝势方法，即把原子核的库仑吸引势加上一个短程的排斥势，两项之和（赝势）在原子核附近变得比较平坦，而得到的本征能量和价电子波函数与真实本征值及在原子核外的真实波函数是一样的。

赝势同时概括了离子实内部的吸引作用和波函数的正交要求，二者相消，如图 8-1c 所示，赝势比离子势弱，比较平坦。对于这样的赝势系统，用平面波展开赝波函数可以很快收敛。实际采用的赝势总是要使价电子波函数在离子实内部尽可能的平坦，且在离子实之间的区域给出与采用真实势相同的波函数，见图 8-1d。值得指出的是，虽然 ψ_V^{ps} 是赝波函数，但由此得到的能量却是相应于晶体真实价电子波函数的本征能量 E_V 。

方程（8-25）中的赝势是有效势，它包含离子实的作用（称为离子赝势或原子赝势）和价电子的作用。在赝势方法中，使用的离子赝势（原子赝势）可分成三类：经验赝势（EPM）、半经验原子赝势和第一性原理从头计算原子赝势。

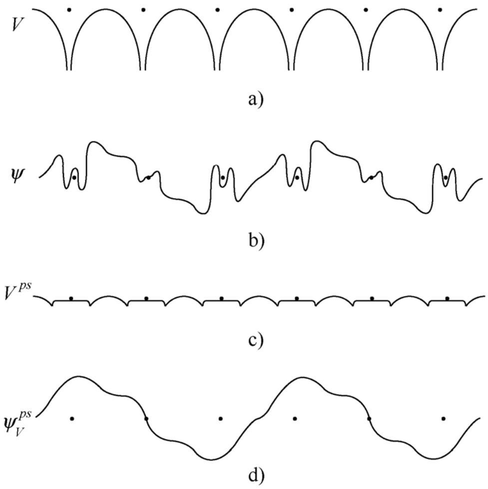


图 8-1　晶体周期势（图 a）、Bloch 波函数（图 b）与赝势（图 c）、赝波函数（图 d）的比较

在经验赝势方法中，晶体势被表示成原子势的叠加，在赝势拟合过程中经过反复与实验数据（一般是能带、态密度、响应函数等）作比较、修改，直至得到与实验数据接近的结果。目前，经验赝势方法最主要的用途是在现代从头算原子赝势自洽迭代计算中做初始值使用。模型赝势是用于自洽计算的半经验原子赝势，在这种赝势表达式中含有一个或几个可变参量，用与实验数据相比较的办法来确定这些参量。没有任何附加经验参数的原子赝势，称为第一性原理从头算原子赝势。目前，在能带理论计算中最常用的从头算原子赝势是 D. R. Hamann 提出的模守恒赝势。

理论计算中最常用的赝势是模守恒赝势以及超软赝势，这些赝势所对应的波函数不仅与真实势对应的波函数具有同样的能量本征值，而且在 $\boldsymbol{r}_c$ 以外，与真实波函数的形状和幅度都相同，另外在 $\boldsymbol{r}_c$ 以内变化缓慢，没有大的动能。

在密度泛函理论的 LDA 和 GGA 的框架下，赝势的首要用途是在现代从头计算原子赝势自洽迭代计算中作初始值使用。这种赝势的最大优点是它的易于处理的简单形式，能以最小的代价用于任意晶体的能带计算。虽然这种经验赝势产生的电荷分布与真实的不完全一致，但它能正确地体现出晶体的对称性，在一个耗费很大的第一性原理赝势自洽从头计算中给出了一个很好的初始值。

赝势应用过程中的一个概念就是赝势的硬度。如果一个赝势使用很小的傅里叶格子就取得精确的结果，则称此种赝势为软的，反之则称之为硬的。现在在能带理论计算中最常用的 Hamann 等提出的模守恒赝势就是一种相当硬的质势。Vanderbilt 所提出来的超软赝势的想法是不用释放非收敛性条件的，用这样的方法来产生更软的赝势。在这种方法里，虚波（seudo−wave）函数在核心范围被允许

做成尽可能的软，以得到尽可能小的截止能量。从技术上而言，这是靠广义的正交条件来实现的。为了要得到总的电子密度，波函数平方所得到的电荷密度必须在核心范围内再加上附加额外的密度。这个电子云密度因此被分成两个部分，第一部分是一个延伸至整个单位晶胞的平滑部分，第二部分是一个局域化在核心区域的自旋部分。前面所提的附加部分只出现在密度里，并不出现在波函数中。超软赝势除了比模守恒赝势更软外，还有另外一个好处。采用超软赝势保证了在预先选择的能量范围内会有良好的散射性质，这使得产生的赝势具有更好的转换性和精确性。

1. 模守恒赝势（norm conserving pseudopotential，NCPP）

除了经验势和可用于自洽计算的半经验模型赝势外，常用的第一性原理赝势方法为模守恒赝势。它要求赝势波函数满足如下条件。

1）本征值与真实本征值相等。

2）没有节点。

3）在原子核区之外（$\boldsymbol{r} > \boldsymbol{r}_c$）与真实波函数相等。

4）在内层区内（$\boldsymbol{r}<\boldsymbol{r}_c$）的赝电荷与真实电荷相等，将赝波函数插入到薛定谔方程中即得对应的赝势。一般说来，小的 $\boldsymbol{r}_c$ 移植性好，可用于不同环境，但平面波收敛慢。

模守恒赝势产生的基本思想是对于孤立原子，构造一个参数化的等效势 V，在给定交换关联模型的条件下，求解的单电子方程得到赝波函数 φ_{ps} 和本征值；在相同交换关联模型下求解全电子方程，得到波函数 ψ 和本征值。调整参数化势能函数 U，使得到的两组本征值相同，在某个截断半径 $\boldsymbol{r}_c$ 以外波函数一致，在该截断半径以内保证波函数的模方一致，此时的等效势 V 有效地描述了价电子感觉到的来自原子核和处于稳定状态的满壳层电子的相互作用，这个等效势就是模守恒赝势 U_{ps}。赝势的构造有可能将波函数急剧振荡的芯区用一个平滑的波函数替代而不改变电荷密度分布。平滑的波函数在用平面波展开过程中需要的平面波矢要小得多，在晶体环境中，这意味着截断能的降低，进而导致收敛性的改善。

第一性原理模守恒赝势可分为局域和非局域两部分

$$U_{ps} = \sum_{v} V_{loc}(\boldsymbol{r} - \boldsymbol{R}_v) + \sum_{v} U_{NL}(\boldsymbol{r} - \boldsymbol{R}_v, \boldsymbol{r}' - \boldsymbol{R}_v) \tag{8-26}$$

其中，v 是对离子势求和。考虑到原子球对称性，得用球谐函数将赝势的非局域部分写成

$$\begin{aligned} U_{NL}(\boldsymbol{r},\boldsymbol{r}') &= \sum_{l,m} Y_{lm}^{*}(\theta',\varphi') Y_{lm}(\theta,\varphi) V_l(\boldsymbol{r},\boldsymbol{r}') \\ &= \sum_{l,m} |lm\rangle\langle lm| V_l(\boldsymbol{r},\boldsymbol{r}') \end{aligned} \tag{8-27}$$

如果将 $V_l(\boldsymbol{r},\boldsymbol{r}')$ 取成半局域形式，即径向是局域的，只有角部分是非局域的：

$V_l(\boldsymbol{r},\boldsymbol{r}')=V_l(\boldsymbol{r})\delta(\boldsymbol{r},\boldsymbol{r}')$，并定义角动量 l 的投影算符 $P_l=\sum_m |lm\rangle\langle lm|$ 。则半局域的原子赝势可以写成如下的形式

$$U_{ps}(\boldsymbol{r})=\sum_v V_{loc}(\boldsymbol{r}-\boldsymbol{R}_v)+\sum_{vl} V_l(\boldsymbol{r}-\boldsymbol{R}_v)\dot{P}_l \tag{8-28}$$

为了简化计算，Kleinman 和 Bylander（KB）将上面半局域赝势部分用一个非局域赝势来近似

$$U_{NL}(\boldsymbol{r},\boldsymbol{r}')=\sum_{l,m}\frac{|V_l\lambda_m\rangle\langle V_l\lambda_m|}{\langle\lambda_m|V_l|\lambda_m\rangle} \tag{8-29}$$

模守恒赝势方法可以在局域密度近似下采用平面波基精确有效地计算固态性质，且可移植性好，但在描述局域价轨平面波基仍然很大，因而在第 I 族元素和过渡金属中的应用受到了限制，通过优化光滑的赝波函数或赝势和增大截断半径 $\boldsymbol{r}_c$ 的改进有一定效果，但模守恒条件的限制使得在一些情况下，如 O 2p 或 Ni 3d 轨道，很难构造出比全电子波函数更光滑的赝波函数，收敛仍然很慢。

模守恒型的赝势从最早的 Hamann 等提出，后来建立一组涵盖整个周期表的参数，之后的 Kerker、TM 赝势，一直到 Optimised 赝势，都是在朝着兼顾准确性的情况下，尽可能使必须使用的平面波基底数目越少越好，它是直接影响着所需计算量大小的量。一个赝势所需的基底数多少，可由系统总能 E_{tot} 对截断动能 E_{cut} 的收敛性来判断，即平面波截断动能 E_{cut} 用到多大时则固态计算所求得系统总能就不再改变，所需 E_{cut} 越小，也就是所谓的赝势越“软”。

使用 Optimised 或 TM 赝势虽然能够把模守恒型赝势变得很“软”，但模守恒条件对于原本就已经没有节点价电子云分布的改造及最佳化的程度，与现今日渐普遍的超软赝势（它不必遵守模守恒条件）来比，节省计算的程度仍是有限的。总之，计算量的大小是取决于原子的种类这一点，是十分明确而普遍的认识，也就是说不同种类元素其势的“软硬”差异会令人明显感觉到。

2. **超软赝势**（ultrasoft pseudopotential，USPP）

事实上，模守恒条件并不是必须的，任何构造一套光滑的赝波函数，只要它能够通过某种方式从赝波函数得到真实的电荷密度，将之插入到具有正确本征值的径向薛定谔方程中得到相应赝势，就能实现计算目的。在此前提下，收敛越快的越好。Vauderbilt 提出的超软赝势方法，它是通过对模守恒条件的弛豫而发展了的一套方法。超软赝势的赝波函数已经不必再遵守模守恒条件，而是通过定义附加电荷来达到所谓的模守恒条件的，其目的基本上是将被砍掉的较局域化的电子云补回去。它构造的赝波函数在内层区之外（$\boldsymbol{r}>\boldsymbol{r}_c$）和全电子波函数一致，而在内层区内（$\boldsymbol{r}<\boldsymbol{r}_c$），通过引入一个广义的正交条件，用于自洽计算的真实电荷密度可以从赝波函数计算得到。图 8-2 给出了氧原子 2p 径向波函数、模守恒赝波函数

和超软赝波函数的比较。

超软赝势目前已经成功地应用于过渡金属和第 I 族元素计算中。

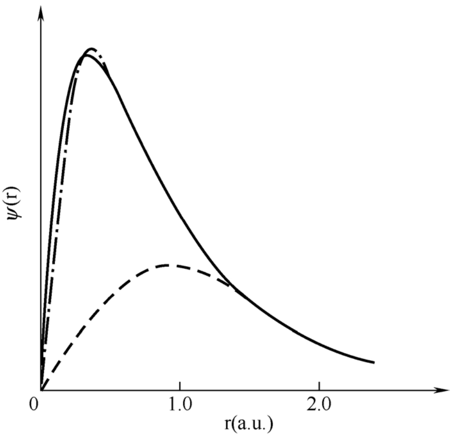


图 8-2　氧原子 2p 径向波函数（实线）、模守恒赝波函数（点画线）、超软赝波函数（虚线）的比较

8.3　第一性原理的计算方法

8.3.1　波函数的展开

在实际求解科恩–沈［吕九］方程时，由于原子核产生的势场项在原子中心是发散的，波函数变化剧烈，需要用大量的平面波展开，因而计算变得十分困难，所以需要尽可能包含少的基函数。好的基应该收敛快，计算中可以包含尽量少的维度。根据不同研究对象，有很多不同的选择基函数的方法，如原子轨道线性组合法（linear combination of atomic orbitals，LCAO）、平面波方法（plane wave，PW）、缀加平面波法（augmented plane wave，APW）、格林函数法（KKR）、线性缀加平面波法（linear augmented plane wave，LAPW）、Muffin-tin 轨道线性组合法（linear muffin-tin orbitals，LMTO）等。

1. 布里渊区积分

在通常的能带图中，经常会出现比如 Γ、Λ、K、X、…符号。这些符号表示的是布里渊区内的高对称性的一些特殊的 ***k*** 点。这些 ***k*** 点有特殊重要的意义，是因为在进行体系总能计算时，通常要对布里渊区内的波函数或本征值进行积分，在实际计算过程中，积分是通过对部分特殊选取的 ***k*** 点的求和完成的。比较常见的 ***k*** 点网格的撒取方法有 Monkhorst-Pack 方法。

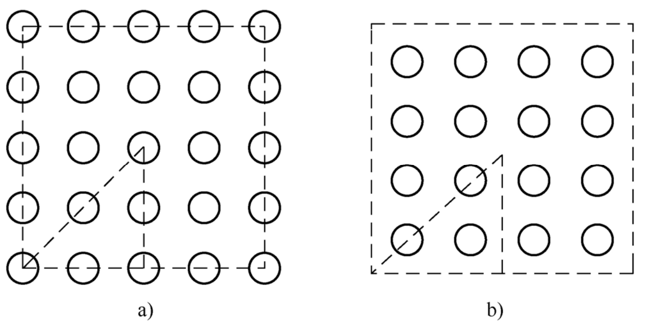
众所周知，根据布洛赫（Bloch）定理，周期体系中的电子波函数可以表述为调幅平面波的形式，即

$$\psi(\boldsymbol{k},\boldsymbol{r}) = \mu(\boldsymbol{r})\exp(i\boldsymbol{k}\cdot\boldsymbol{r}) \tag{8-30}$$

其中，$\boldsymbol{k}$ 是电子波矢，$\mu(\boldsymbol{r})$是具有晶体平移周期性的周期函数。其本征能量和本征矢量为：$E_n(\boldsymbol{k})$，$\Phi_n(\boldsymbol{k})$。不同的电子状态按照量子数 $\boldsymbol{n}$ 进行分类，而量子数 $\boldsymbol{n}$ 则表征能态的分立性。研究多体体系的价电子问题，归根结底是计算出不同类的电子状态的本征值和本征矢量，体系处于基态情况下，哪些不同 $\boldsymbol{k}$ 的低能量状态被电子占据。因体系具有周期性，所以，第一布里渊区的所有 $\boldsymbol{k}$ 可以代表所有的 $\boldsymbol{k}$。但是，由周期边界条件确定的 $\boldsymbol{k}$ 有无穷多个，而且计及相互作用势的实际体系中许可的 $\boldsymbol{k}$ 在倒易空间内是不均匀的，实际计算过程中只能选取有限个点。

对周期性体系，电子结构计算与分析都涉及布里渊区（BZ）积分，通过玻恩−冯卡门边界条件，在实际计算中积分变成求和的形式。更一般地，积分可以变成对部分特殊选取的 $\boldsymbol{k}$ 点进行带权重的相加。这样，就涉及到 $\boldsymbol{k}$ 点的选取。在计算过程中，要用尽量少的 $\boldsymbol{k}$ 点得到尽可能精确的积分。对任意函数，如果取如上两个 $\boldsymbol{k}$ 点，可以保证其傅里叶展开以后直到 2 阶分量的积分为零。因此，为了获得较高的效率需要选取一些特殊的 $\boldsymbol{k}$ 点，原则是对一定的 $\boldsymbol{k}$ 点数量使得尽可能高的傅里叶分量的积分都为零。同时，$\boldsymbol{k}$ 点的选取也需简便易行，综合考虑后较好的方法为 Monkhorst-Pack（MP）取点法，它将三维问题转化为三个一维问题。一套 MP 网格是 BZ 内均匀分布的规则格点集，其数目由参数 $n_{\mathrm{k1}}\times n_{\mathrm{k2}}\times n_{\mathrm{k3}}\,n$ 决定。MP 的 $\boldsymbol{k}$ 点的分数坐标取为 $u_r = (2r - n_k - 1)/2\,n_k \quad r = 1,2,3,\cdots,n_k$ 。

Monkhorst-Pack 方法的 $\boldsymbol{k}$-grid 对偶数分割不包括原点，这样可减少需要计算的点的数量，如图 8-3 所示。经验表明，对 $n_{\mathrm{k}}\leqslant 8$，应尽量采取偶数分割，否则采取奇数分割。对六角晶格，因为晶格矢量不垂直，MP 分数坐标公式需作一定的修正，此时应该将 Γ 点包含在 $\boldsymbol{k}$ 网格中。与绝缘体比较，金属需要更密的 $\boldsymbol{k}$ 点。



a)　　　　b)

图 8-3　Monkhorst-Pack 方法的 $\boldsymbol{k}$ 点选择

a）含原点，需计算 6 个点；b）不含原点，只需计算 3 个点

2. 平面波（PW）方法

平面波方法是以平面波为基函数展开电子波函数。这种方法的数学形式相当简单。在大多数情况下，哈密顿量矩阵元在平面波基下可以简单地用解析式表达，而且平面波基是非定域的，不依赖于原子的位置。但晶体的电子波函数占有很宽的动量范围，在靠近原子核的区域，电子动量很大，波函数快速振荡，在远离原子核的区域，电子动量很小，波函数变化平缓。因此，要准确地描述电子运动状态，波函数展开所需的平面波数相当大，使得计算量庞大。为克服这一困难，通常采用动量较小的平面波和具有较大动量的孤立原子芯态波函数构成正交化平面波，用正交化平面波为基函数展开电子波函数，只需要少数几项就可以很好地描述晶体中的电子运动状态。

平面波是最简单的正交完备基，根据布洛赫定理，单电子波函数可以用平面波展开

$$\psi_{nk}=\sum_{G}c_n(\boldsymbol{K}+\boldsymbol{G})\exp[i(\boldsymbol{K}+\boldsymbol{G})\bullet\boldsymbol{r}] \tag{8-31}$$

式中，$\boldsymbol{G}$ 是原胞的倒格矢，$\boldsymbol{K}$ 是第一布里渊区的波矢，$c_n(\boldsymbol{K}+\boldsymbol{G})$是平面波展开系数（简称平面波系数）。在对真实系统的模拟中，由于电子数目的无限性，$\boldsymbol{K}$ 矢量的个数从原则上讲应该是无限的。但由于 $\psi(\boldsymbol{r})$随 $\boldsymbol{K}$ 点的变化在 $\boldsymbol{K}$ 点附近是可以忽略的，因此可以使用 $\boldsymbol{K}$ 点取样通过有限个 $\boldsymbol{K}$ 点进行计算。另一方面，为了得到对波函数的准确表示，$\boldsymbol{G}$ 矢量的个数也应该是无限的。但由于对有限个数的 $\boldsymbol{G}$ 矢量求和已经能够达到足够的准确性，因此对 $\boldsymbol{G}$ 的求和可以截断成有限的，给定一个截止能

$$E_{\rm cut}=\frac{\hbar(\boldsymbol{G}+\boldsymbol{K})^2}{2m} \tag{8-32}$$

对 $\boldsymbol{G}$ 的求和可以限制在$(\boldsymbol{G}+\boldsymbol{K})\leqslant E_{\rm cut}$的范围内，即要求用于展开的波函数的能量小于 $E_{\rm cut}$。当 $\boldsymbol{K}=0$ 时，即在 Γ 点，有很大的计算优势，因为这时波函数的相因子是任意的，就可以取实的单粒子轨道波函数。

哈密顿矩阵元为

$$H_{\boldsymbol{GG'}}=\frac{1}{2}\left|\boldsymbol{K}+\boldsymbol{G}\right|^2\delta_{\boldsymbol{GG'}}+V_{eff}(\boldsymbol{G}-\boldsymbol{G'}) \tag{8-33}$$

原则上无穷多个平面波才构成一个完备集，但具有较小动能$|\boldsymbol{K}+\boldsymbol{G}|^2/2$ 的平面波系数 $c_n(\boldsymbol{K}+\boldsymbol{G})$比具有较大动能的平面波系数大，因此可以只用小于某一能量 $E_{\rm cut}$ 的平面波作为基进行展开。$E_{\rm cut}$ 越小，计算越容易，但截断所引起的误差也越大，因而需要增加 $E_{\rm cut}$ 直到收敛。

由于固体中离子实附近具有很强的局域势，在远离原子核处，波函数的行为趋近于一个平面波，但在近核处变化很快，全电子计算的平面波展开收敛太慢。众所周知，固体中许多物理性质主要由费米面附近的价电子决定，赝势方法利用

这个性质，采用一个减弱的赝势来代替内层电子和原子核势场，相应的赝势作用下的波函数不再有振荡特征，而是平滑的赝波函，因此大大减少了平面波展开所需要的平面波函数目。

可将真实价波函数 $\psi_n(\boldsymbol{k}, \boldsymbol{r})$看做是由赝势波函数 $\lambda_n(\boldsymbol{k},\boldsymbol{r})$和内层波函数 $\phi_J(\boldsymbol{k},\boldsymbol{r})$ 线性组合

$$\psi_n(\boldsymbol{k},\boldsymbol{r}) = \lambda_n(\boldsymbol{k},\boldsymbol{r}) - \sum_J \sigma_{nJ}(\boldsymbol{k})\,\phi_J(\boldsymbol{k},\boldsymbol{r}) \tag{8-34}$$

其中系数 $\sigma_{nJ}(\boldsymbol{k})$ 由正交条件 $\int \mathrm{d}\boldsymbol{r}'\phi_J^*(\boldsymbol{k},\boldsymbol{r}')\,\psi_n(\boldsymbol{k},\boldsymbol{r}') = 0$ 确定

$$\sigma_{nJ}(\boldsymbol{k}) = \int \mathrm{d}\boldsymbol{r}'\phi_J^*(\boldsymbol{k},\boldsymbol{r}')\lambda_n(\boldsymbol{k},\boldsymbol{r}') \tag{8-35}$$

联合真实波函数 $\psi_n(\boldsymbol{k},\boldsymbol{r})$所满足的薛定谔方程 $[T + V(\boldsymbol{r})]\psi_n(\boldsymbol{k},\boldsymbol{r}) = E_n(\boldsymbol{k})\psi_n(\boldsymbol{k},\boldsymbol{r})$，可得到赝波函数满足如下方程

$$[T + U_{ps}]\lambda_n(\boldsymbol{k},\boldsymbol{r}) = E_n(\boldsymbol{k})\lambda_n(\boldsymbol{k},\boldsymbol{r}) \tag{8-36}$$

$$U_{ps}\lambda_n(\boldsymbol{k},\boldsymbol{r}) = V(\boldsymbol{r})\lambda_n(\boldsymbol{k},\boldsymbol{r}) + \int \mathrm{d}\boldsymbol{r}' V_R(\boldsymbol{r},\boldsymbol{r}')\lambda_n(\boldsymbol{k},\boldsymbol{r}') \tag{8-37}$$

$$V_R(\boldsymbol{r},\boldsymbol{r}') = \sum_J \phi_J^*(\boldsymbol{k},\boldsymbol{r}')[E(\boldsymbol{k}) - E_J]\phi_J(\boldsymbol{k},\boldsymbol{r}') \tag{8-38}$$

上式中 U_{ps} 称为原子赝势。根据密度泛函理论，原子赝势包括离子赝势 U_{ps}^{ion}、价电子库仑势和交换关联势：$U_{ps} = U_{ps}^{ion} + V_H^{ps}(\boldsymbol{r}) + V_{xc}(\boldsymbol{r})$。其中后两项 $V_H^{ps}(\boldsymbol{r})$ 和 $V_{xc}(\boldsymbol{r})$ 可以从真实电荷密度计算（此时等于对应的全电子势 V_H 和 V_{xc}），也可从赝电荷密度计算。

从上面可知：

1）赝波函数和真实波函数具有完全相同的能量本征值 $E_n[\boldsymbol{k}]$，这是赝势方法的重要特点。

2）赝势第二项是排斥势，与真实的吸引势有相消趋势，因此比真实势弱。

3）赝势包括局域项，其中非局域项同时与 $\boldsymbol{r}$ 和 $\boldsymbol{r}'$处的赝波函数 $\lambda_n(\boldsymbol{k}, \boldsymbol{r})$和 $\lambda_n(\boldsymbol{k}, \boldsymbol{r}')$有关，而且依赖于其能量本征值 $E_n(\boldsymbol{k})$。

利用平面波基来计算有很多优点。首先平面波基能很方便地采用快速傅里叶变换（FFT）技术，使能量、力等的计算在实空间和倒空间快速转换，这样计算尽可能在方便的空间中进行。如前面讲到的哈密顿量中的动能项的矩阵元，在倒空间中只有对角元非零，就比实空间减少了工作量。第二，平面波基函数的具体形式并不依赖于核的坐标。这样，一方面，价电子对离子的作用力可以直接应用赫尔曼−费恩曼（Hellman−Feymann）定理得到解析表达式，计算显得非常方便。另一方面也使总能量的计算在不同的原子构型下有基本相同的精度。此外，平面波计算的收敛性和精确性比较容易控制，因为通过截断能 E_{cut} 的选择可以方便地改变平面波基的多少。当然平面波基也有缺点，一般电子轨道具有一定的局域性，

而平面波是空间均匀的，因此电子轨道展开时与原子轨道基相比，平面波基的个数要多得多。为了尽量减少平面波基的个数，一般在平面波的计算中都采用赝势来描述离子实与价电子之间的相互作用，使电子轨道波函数在离子实内部的分布尽量平缓些。

3. LCAO 方法

原子轨函线性组合（LCAO）方法的中心思想就是用原子轨道的线性组合来作为一组基函数，由此求解单电子薛定谔方程。这种方法在绝缘体的能带结构研究中是成功的。处于同一格点上的原子轨道是正交的，但不同格点上的原子轨道函数一般是非正交的，由它们组成的基函数也是非正交的，在计算中必然会遇到多中心积分的计算问题，计算量相当庞大。为了克服这些困难，人们提出了许多方案，如 Slater−Koster 参数方法和 W.A.Harrison 键轨道近似。

Slater−Koster 参数方法将原子轨道线性组合方法中的哈密顿量矩阵元看成参数，其大小由布里渊区中心或边界上高对称性电子波矢的精确理论值或实验值拟合得到。一般来说 Slater−Koster 参数方法对半导体价带的描述较好，但对导带的描述却不太好。为了得到对价带和导带较好的描述，往往大量引入参数，有时参数达到 50 个之多。价键轨道近似在化学中应用广泛，但应用于固体时，由于忽略了成键态与反键态的耦合作用，在半导体中，对类 s 价带的描述较好，但对价带顶的类 p 价带的描述却不好，没有色散，反映在电子态密度上为一个 δ 函数。为改善对类 p 价带的描述，W.A.Harrison 还引进了额外的相互作用参数。此外键轨道近似对导带的处理也不理想。

4. APW 方法

缀加平面波方法和糕模（muffin−tin）轨道线性组合方法有一个共同特点，即它们都与糕模（muffin−tin）势有关。muffin−tin 势将晶体空间分为两个区域：以每个原子为中心的 muffin−tin 球形区以及球外的球间区域。假定在球内 muffin−tin 势具有球对称性，球间区域则常取为常数势。muffin−tin 势通常被用于金属与合金的电子结构计算。缀加平面波方法可以给出整个能带结构的很精确的结果，但由于要求解的久期方程是一个超越方程，计算量还是很大的。原子球近似下的 muffin−tin 轨道线性组合方法在对分子、以四面体结构结合的绝缘体和半导体的单电子能量和结合能计算中会出现严重误差。而对于这种方法加以非 muffin−tin 近似修正是十分复杂的。

8.3.2　自洽场（SCF）运算

在实际的计算中，科恩−沈［吕九］方程是通过自洽计算来求解的，其计算流程如图 8-4 所示。图中的流程为：将多体系统原胞划分为足够细的网格点，在每个网格点上初始化一组试探波函数（通常设为随机数），然后算出网格上的科恩−沈［吕

九］势，求解本征方程。解出来的本征函数的值与初始化的试探波函数的值一般不会相同，再将新解出来的波函数的一部分叠加到初始值上重新计算科恩–沈［吕九］势，利用修正过的势再次求解本征方程。所得到的本征函数又用于修正上一步循环输入的波函数……循环迭代的结果，是最终求得的本征函数的值不再变化，计算得以收敛。利用收敛后的这组单电子波函数，就可得到体系总能量和电荷密度分布。

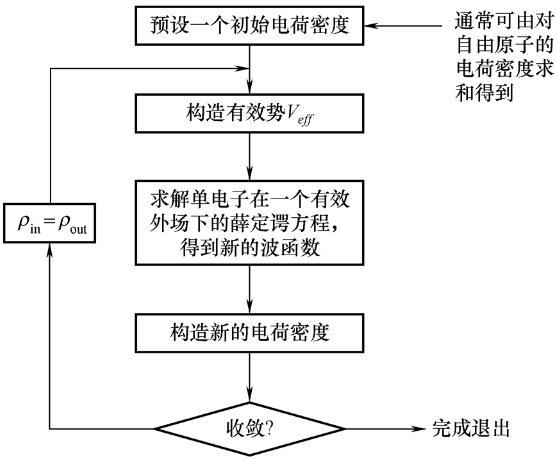


图 8-4　自洽方法求解科恩–沈［吕九］方程流程图

1．求解单电子波函数

求解单电子波函数本质上是一个求本征值问题。对小矩阵，有很多数值算法可以直接求解。然而，电子结构计算中通常遇到的是很大的矩阵，无法直接对角化。当只对少数最低的本征值感兴趣时，可以用迭代算法求解。常见的迭代算法有，单带共轭梯度法、Davidson 法与 RMM-DIIS 法以及 Lanczos 法，这里不再详细叙述，请参阅其他有关资料。

2．求电荷密度与能量

（1）分数占据（Partial Occupation）

在前面讨论布里渊区积分时，没有考虑占据数（零温下为阶梯函数）的影响。

$$\sum_n \Omega_{BZ}\int_{BZ}\left\langle \Phi_n(\boldsymbol{k})\middle|\hat{X}\middle|\Phi_n(\boldsymbol{k})\right\rangle \theta\left(\varepsilon_{nk}-\mu\right)\mathrm{d}\boldsymbol{k} \tag{8-39}$$

事实上，阶梯函数在数值上不稳定，使得布里渊区积分随 $\boldsymbol{k}$ 点数目收敛很慢。此时，常引入分数占据来代替阶梯函数。

1）费米–狄拉克 Fermi-Dirac 分布函数

$$f(\varepsilon)=\frac{1}{\mathrm{e}^{\frac{\varepsilon-\mu}{k_B T}}+1} \tag{8-40}$$

优点：温度具有物理意义。

2）高斯（Gauss）型

$$f(\varepsilon)=\frac{1}{2}\left[1-\mathrm{erf}\left(\frac{\varepsilon-\mu}{k_B T}\right)\right] \tag{8-41}$$

优点：可以外推得到 T=0 时的能量。

3）Methfessel & Paxton 方法。MP 方法通过将阶梯函数用一组正交函数展开来实现，Gaussian 函数是其零阶近似。因此，我们通常说的 MP 方法是指大于 0 阶的情形。对于合理的 $\sigma = k_B T$ 的值，MP 方法中 F 里的熵贡献较小，这样计算的力会更准。一个合理的判据是选择合适的 σ，使得熵小于 1meV/（atom·K）。MP 方法的弱点是不适用于半导体或绝缘体，有可能出现占据数大于 1 的情况。

4）四面体（Tetrahedron）方法

数占据从算法上也可以看成对布里渊区积分离散化时，对不同 $\boldsymbol{k}$ 点的权重进行系统的调整。从这个思路出发，在由四个 $\boldsymbol{k}$ 点组成四面体内做线性插值可以得到修正的权重，也即等效的分数占据函数 f_{nk}。实际应用上常采用 Blochl 等人提出的修正的四面体算法。

1）自动选择四面体使误差尽量抵消。

2）对金属，误差抵消不完全，加入修正项。

3）最后积分形式仍为加权求和。

修正的四面体算法可以十分准确地计算能量、态密度等量，且不需要输入额外的参数。但是，由于它对分数占据不满足变分原理，所以得到的力可能不够准确，不能用来优化金属体系。

（2）能量泛函

在 SCF 迭代过程中，运用类似与热力学中的勒让德（Legendre）变换，能量泛函可以写成不同的形式。收敛以后它们都得到同一个能量，但是它们趋向于收敛的行为各不相同。KS 能量泛函的原始定义为

$$\begin{aligned} E_{KS} &= T_S[\rho]+E_{pot}[\rho] \\ &= E_S-\int \mathrm{d}\boldsymbol{r} v^{in}(\boldsymbol{r})\rho^{out}(\boldsymbol{r})+\int \mathrm{d}\boldsymbol{r} v_{ext}(\boldsymbol{r})n(\boldsymbol{r})+E_H[\rho]+E_{XC}[\rho] \end{aligned} \tag{8-42}$$

KS 能量泛函原则上是电荷密度的泛函，但是在实际操作中可以看成有效势的泛函

$$E_{KS}[v^{in}]=E_S[v^{in}]-\int \mathrm{d}\boldsymbol{r} v^{in}(\boldsymbol{r})\rho^{out}(\boldsymbol{r})+E_{pot}[\rho^{out}] \tag{8-43}$$

它满足变分原理。我们也选择另外一种显式的密度泛函形式，这就是常用的 Harris-Foulkes（HF）泛函。因为输入势可以由输入密度决定，我们将 $v[\rho^{in}]$ 记为 $v_{\rho^{in}}$。

$$E_{HF}[\rho^{in}] = E_S[v_{\rho^{in}}] - \int \mathrm{d}\boldsymbol{r} v^{in}(\boldsymbol{r})\rho^{in}(\boldsymbol{r}) + E_{pot}[\rho^{in}] \tag{8-44}$$

因为自洽收敛时 $\rho^{in} = \rho^{out}$，*HF* 泛函显然与 *KS* 泛函有同样的收敛性质。但是，*HF* 泛函不满足变分原理，它主要用在非自洽计算（无需计算 ρ^{out}）时的能量估计，而且在自洽计算的过程中 *HF* 泛函给出的能量值也通常比 *KS* 泛函更接近真实能量。

3. SCF 迭代：mixing

SCF 迭代刚开始时 ρ^{in} 和 ρ^{out} 可能会相差很大，使得整个算法不稳定，收敛困难。这时需要做一定的 mixing。

（1）Linear mixing

$$\begin{aligned}\rho_{i+1}^{in} &= \alpha\rho_i^{out} + (1-\alpha)\rho_i^{in} \\ &= \rho_i^{in} + \alpha(\rho_i^{out} - \rho_i^{in})\end{aligned} \tag{8-45}$$

为了考察的 α 取法，定义与真实密度的差为 $\delta\rho = \rho - \rho_{KS}$。在基态密度附近线性化，$\delta\rho^{out} = (\tilde{\chi}+1)\delta\rho^{in}$，若 $\rho_{i+1}^{in} = \rho_{KS}$，则可收敛。可以证明，欲使 SCF 收敛，则有：$\alpha < \dfrac{2}{\tilde{\chi}_{\max}^{-1}} = 2\tilde{\chi}_{\min}$。

从物理上来看，响应函数可以看成极化率的度量。所以，对束缚很强的刚性系统，可以取较大的 α，而对于金属表面之类的较软的系统，则比较困难。

（2）其他 Mixing 算法

对介电函数作均匀电子气 Thomas-Fermi 屏蔽近似，可以得到 Kerker Mixing，其形式与 linear mixing 基本相似，这里不作具体介绍。

8.3.3 结构优化

对于给定各原子位置、元素种类的体系，通过密度泛函理论自洽求解科恩-沈［吕九］方程可以得到整个系统处于多电子基态时的总能。总能量对系统虚拟微位移的导数就是各原子的受力（Hellmann-Feynman）。这为理论上预言物质的结构提供了一种行之有效的方法。因为自然界稳定的结构应该具有最低的总能，我们只要根据原子受力来变化原子的位置，直到整个体系的总能达到最低（所有原子受力为零，当然在实际的计算过程中，我们只能给出希望达到而且有限的计算精度），即找到能量面的（全局）最小值，这时所对应的物质结构就是自然界最稳定的结构。该过程被称为结构优化。

结构优化是通过调节结构模型的几何参数来获得稳定结构的过程，其结果是使模型结构尽可能地接近真实结构。进行结构优化的判据可以根据研究的需要而定，一般是几个判据组合使用。常用的判据有以下几个。

1）自洽场收敛判据：对给定的结构模型进行自洽场计算时，相继两次自洽

计算得到的晶体总能量之差足够小，即相继两次自洽计算的晶体总能量之差小于设定的最大值。

2）力判据：每个原子所受的晶体内作用力（Hellmann-Feynman）足够小，即单个原子受力小于设定的最大值。

3）应力判据：每个结构模型单元中的应力足够小，即应力小于设定的最大值。

4）位移判据：相继两次结构参数变化引起的原子位移的分量足够小，即原子位移的分量小于设定的最大值。

为了确保搜索能量面的最小值时能找到全局最小而不是局域最小，并提高整个搜索过程的效率，我们需要一些强有力的搜索算法以使原子最快地运动到最稳定结构的位置。最常用的方法有直接能量最小化、最深梯度（即最大受力）法、共轭梯度法（考虑到前后两步的受力是否为同一方向）、准牛顿方法、阻尼动力学法等。这里主要介绍材料模拟常用的几种优化方法。

1. 最速下降法

最速下降法是沿着局部净受力方向行走的，以进行能量极小。从初始点开始，沿着局部梯度的反方向 $-g_l$，并通过在此方向上的一维极小化，移动到该方向的极小点，再从这个点开始重复以上过程，直到达到所要求的精度。相邻两步的梯度（因而相邻两步的运动方向）正交。最速下降法在远离极小点效率很好，在接近极小时效率不高，而且沿梯度方向每前进一步将对接下来一步都引入一个正比于它梯度的误差，常常只在优化的最初几步使用这种方法。

2. 共轭梯度法

共轭梯度法克服了最速下降法的困难。在此方法中，每相邻两次优化的起始点的 g_l 仍是正交的，但优化方向 v_l 由当前梯度 g_l 结合前一次优化方向 v_{l-1} 和梯度 g_{l-1} 共同决定，v_l 与 v_{l-1} 互为共轭

$$v_l = g_l + \gamma_l d_{l-1} \tag{8-46}$$

其中 γ_l 是一个标量数，由一次前优化起点的梯度 g_{l-1} 和优化终点的梯度（即当前时刻梯度）g_l 共同决定，不同的算法给出各自的确定公式。

Flrtcher-Reeves 算法：

$$\gamma_l = \frac{g_l \bullet g_l}{g_{l-1} \bullet g_{l-1}} \tag{8-47}$$

Folak-Ribiete 算法：

$$\gamma_l = \frac{(\overline{g}_l - g_{l-1}) \bullet g_{l-1}}{g_{l-1} \bullet g_{l-1}} \tag{8-48}$$

对于能量函数 $f(P)$，我们可以按如下方式进行优化：

若初始点在 P_0，令 $v_0 = -g_0 = -\nabla f(P_0)$，沿 v_0 方向运用一维极小化方法到达该方向的一极小点 P_1，则 $g_1 = \nabla f(P_1)$。由 g_0 和 g_1 可得到 γ_1，则 $v_1 = -g_0 + \gamma_0 v_0$，再沿

v_1 找极小，重复以上过程，如果函数是含 N 个变量的二次型，则通过 N 次一维极小化就可以找到极小。

上面两种方法在优化中只用了势能函数的一阶导数，即梯度。

8.3.4 能带计算

固体能带理论是研究固体中电子运动的一个重要理论，是固体电子论的支柱。建立布洛赫（Bloch）定理之上的固体能带理论用量子力学方法来确定固体电子能级（能带），并用以阐明和解释固体的许多基本性质，如电导率、热导率、磁有序、光学介电函数、振动谱等。本节主要介绍能带理论的理论基础、基本概念、能带计算方法以及应用能带理论对固体物理性质的计算。

能带理论是一个近似理论。固体中有大量的原子核和电子，严格求解这种多粒子系统的薛定谔方程是不可能的，必须采用一些近似和简化。能带理论采用了如下三个近似：

1）玻恩−奥本海默 Born-Oppenheimer 近似（绝热近似），将原子核的运动与电子的运动分开。

2）单电子近似，把每个电子的运动看成是独立的在一个等效势场中的运动，这个等效势场包括原子核的势场和其他电子对该电子的平均作用势（库仑势和交换相关势）。

3）周期性等效势场近似，把固体抽象成具有平移周期性的理想晶体，将固体中电子的运动归结为单电子在周期性势场中的运动。

固体能带计算就是求解对固体系统作了三个近似后得到的单电子方程，前边已经讲过，对这样一个单电子方程进行精确求解是不可能的，只能求近似解。固体能带计算方法很多，有平面波方法、紧束缚近似法、正交平面波法、赝势法、等等。这种采用从头计算法求解固体的单电子薛定谔方程或科恩−沈［吕九］方程得到能带 $E_n(\boldsymbol{k})$和波函数的方法，称为第一性原理方法。值得注意的是，通过周期性势场的近似，固体已被近似成具有周期性结构的理想晶体，将固体能带计算转化为晶体能带计算。

因为能带的计算与前面的内容有所重复，不再另作介绍，可参考前面的章节。

晶体能带及晶体物理性质的计算过程如图 8-5 所示，晶体是一个具有周期性结构的体系，输入时只能给出一个体积有限的晶体结构模型，需利用周期性边界条件，才能得到整个晶体的能带结构。

在选择了计算任务之后，需对计算参数进行设置，如自洽场计算的精度、基组的大小、$\boldsymbol{k}$ 的取值等。自洽场计算的收敛用电子数密度或晶体总能量的收敛来标志，自洽场计算的精度是指收敛的标准。在晶体能带计算中，基组是指用于线性组合成单电子波函数的基函数集合，如平面波方法中的平面波、紧束缚近似方法

中的原子轨道组合成的布洛赫（Bloch）函数。在平面波赝势方法中，给出能量的截断值，确定基组中平面波的数目。能带 $E_n(\boldsymbol{k})$函数是 $\boldsymbol{k}$ 的准连续函数，$\boldsymbol{k}$ 在 $\boldsymbol{k}$ 空间均匀分布，其取值限定在布里渊区中，$\boldsymbol{k}$ 可以取大量的值。由于计算量的原因，计算时 $\boldsymbol{k}$ 只取简约布里渊区中有限的值。在其他条件不变的情况下，$\boldsymbol{k}$ 的取值数目增加，得到的 $E_n(\boldsymbol{k})$函数的精确度增大，但计算量显著增加。

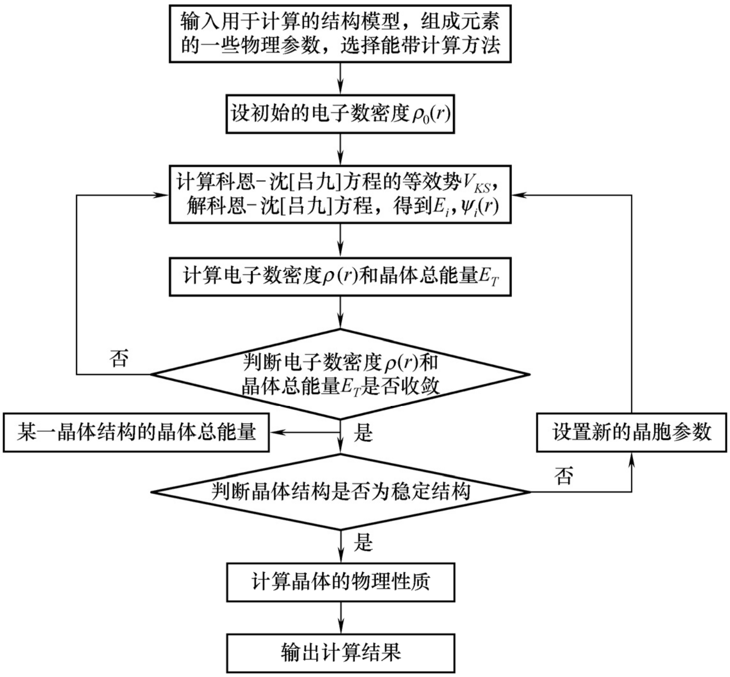


图 8-5　晶体能带及晶体物理性质的计算过程

8.3.5　晶体物理性质的计算

利用自洽场方法求解科恩-沈［吕九］方程，得到所设结构的晶体总能量（单点能）。在计算晶体的物理性质之前必须对所设晶体结构模型进行几何优化，根据关于能量、力、应力、位移的判据来判断晶体结构是否为稳定结构（总能量最小）。如果晶体结构不是稳定结构，重新设置晶格参数进行计算，直至得到稳定的晶体结构。对结构优化后的晶体进行物理性质计算，最后输出计算结果。

1. 晶体的总能量

晶体总能量（不包括核的动能部分）可分成两部分：一部分是原子核与内层电子组成的离子实的能量，这部分能量基本上与晶体结构无关，是一个常数，赝势方法中常把总能量中这部分不变的能量设为零；另一部分是总能量与离子实能

量之差，包括离子实与价电子的相互作用、离子实之间的相互作用以及价电子间的相互作用。晶体总能量与核动能（通常是零点振动能）之和与全部组分原子的孤立原子能之和的差，称为晶体的结合能。

在密度泛函理论中，晶体总能量 E_T 是晶格电子的能量与离子实的排斥能之和

$$E[\rho]=T[\rho]+E_{ext}+E_{cout}+E_{xc}+E_{N-N} \tag{8-49}$$

各项的物理意义与方程（8-11）和方程（8-12）中的相同。其中电子与外场 $v(\boldsymbol{r})$的相互作用能为

$$E_{ext}=\int \mathrm{d}\boldsymbol{r}\,\rho(\boldsymbol{r})v(\boldsymbol{r}) \tag{8-50}$$

电子间库仑相互作用能为

$$E_{coutl}=\frac{1}{2}\iint \mathrm{d}\boldsymbol{r}\mathrm{d}\boldsymbol{r}'\frac{\rho(\boldsymbol{r})\rho(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|} \tag{8-51}$$

在局域密度近似条件下，电子的交换关联能为

$$E_{xc}=\int \mathrm{d}\boldsymbol{r}\,\rho(\boldsymbol{r})\varepsilon_{xc}[\rho(\boldsymbol{r})] \tag{8-52}$$

离子实之间的库仑相互作用解为

$$E_{N-N}=\frac{1}{2}\sum_{\boldsymbol{R},s}\sum_{\boldsymbol{R}',s'}\frac{Z_s Z_{s'}}{|\boldsymbol{R}+\tau_s-\boldsymbol{R}'-\tau_{s'}|} \tag{8-53}$$

这里 Z_s 表示原子 s 的价电子数，$\boldsymbol{R}$ 表示晶格平移矢量，τ_s 表示原胞内原子 s 的相对位矢。动能泛函可通过科恩−沈［吕九］方程用单电子能量表示为

$$T[\rho]=\sum_i\int \mathrm{d}\boldsymbol{r}\,\varPhi_i^*(\boldsymbol{r})(E_i-V_{KS})\varPhi_i(\boldsymbol{r}) \tag{8-54}$$

这里 E_i 是科恩−沈［吕九］方程的本征值。将式（8-50）～式（8-52）、式（8-54）和式（8-19）代入式（8-49），得到晶体总能量

$$E_T=\sum_i E_i-\frac{1}{2}\iint \mathrm{d}\,\boldsymbol{r}\,\mathrm{d}\,\boldsymbol{r}'\frac{\rho(\boldsymbol{r})\rho(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}+\int \mathrm{d}\boldsymbol{r}\,\rho(\boldsymbol{r})\left\{\varepsilon_{xc}[\rho(\boldsymbol{r})]-V_{xc}[\rho(\boldsymbol{r})]\right\}+E_{N-N} \tag{8-55}$$

晶体的单电子能 E 与晶体体积 $N\Omega$ 之间的关系可用 Murnaghan 状态方程

$$E(N\Omega)=\frac{B_0 N\Omega}{B_0'(B_0'-1)}\left[B_0'\left(1-\frac{\Omega_0}{\Omega}\right)+\left(\frac{\Omega_0}{\Omega}\right)^{B_0'}\right]+E(N\Omega_0) \tag{8-56}$$

来描述。方程（8-56）中 N 是晶体中原胞的数目，Ω 是晶格常数尝试值为 a 时的原胞体积，B_0 是体弹性模量，B_0' 是 B_0 对 Ω 的导数。对于不同的晶格常数尝试值 a，即不同的晶体体积 $N\Omega$，可计算出相应的单点能 $E(N\Omega)$，用最小二乘法拟合 Murnaghan 状态方程，便可得到相应于该结构的晶体常数 a_0、体弹性模量 B_0 和该结构的能量极小值 $E(N\Omega_0)$。

对于组成元素确定的体系，可能存在不同的晶体结构。分别对不同的晶体结构进行总能量计算和 Murnaghan 方程拟合，得到相应的能量极小值 $E(N\Omega_0)$，比较

不同晶体结构的能量极小值，便可确定稳定的晶体结构。

由晶体总能量可以确定晶体的一些力学性质，如赫尔曼–费恩曼（Hellmann–Feynman）力和应力。作用于原子 s 上的力 F_s 可由晶体总能量 E_T 对原子 s 的原胞内位矢 $\boldsymbol{\tau}_s$ 求负梯度得到，即 $F_s = -\nabla_{\tau_s} E_T$，这个作用于原子上的晶体内力称为赫尔曼–费恩曼力。

对于晶格应变张量 $\varepsilon = \left|\varepsilon_{\alpha\beta}\right|$，平均应力张量为 $\sigma = \left|\sigma_{\alpha\beta}\right|$，则 $\sigma_{\alpha\beta} = \dfrac{1}{\Omega}\dfrac{\partial E_T}{\partial \varepsilon_{\alpha\beta}}$。

2. 弹性常数

不同晶系的晶体具有不同数量的独立弹性常数。立方结构的晶体存在 3 个独立的弹性常数（C_{11}，C_{12} 和 C_{44}），六方晶系具有 5 个独立的弹性常数（C_{11}，C_{12}，C_{13}，C_{33} 和 C_{44}），而正交晶系则具有 9 个独立性常数（C_{11}，C_{22}，C_{33}，C_{12}，C_{13}，C_{23}，C_{44}，C_{55} 和 C_{66}）。材料的弹性常数描述了它对所加应力的响应，或者反过来说，弹性常数描述了为维持一个给定的形变所需的应力。

为了计算晶体的弹性常数，可先对晶胞进行不同的弹性变形，然后算出变形后的能量，并通过此能量与未变形晶胞的能量差，求出弹性应变能，进而利用弹性常数与应变能的关系计算得到弹性常数。弹性应变能计算式为

$$U = \frac{\Delta E}{V_0} = \frac{1}{2}\sum_{i}^{6}\sum_{i}^{6} C_{ij} e_i e_j \tag{8-57}$$

其中 ΔE 为变形前后晶胞的能量差值；V_0 为原始晶胞的体积；C_{ij} 为弹性常数；e_i 和 e_j 为应变。

8.4　常用的第一性原理计算软件

用于计算材料电子结构的软件有很多，如 CASTEP、WIEN、ABINITI、GAUSSIAN 以及 VASP 等，这些软件的功能大同小异，大多以密度泛函理论为基础，采用不同的波函数展开方法，计算材料的波函数和能带结构，并计算材料系统的总能量，分析材料的各种物理性质。

8.4.1　CASTEP 软件包

CASTEP（Cambridge Sequential Total Energy Package 的缩写）为一个商业软件包，它始于剑桥大学卡文迪许实验室凝聚态理论研究组开发的一系列基于密度泛函方法的第一性原理计算的量子力学程序。CASTEP 属于 Accelrys 公司开发的商业软件 Materials Studio 软件中的一个用来计算和分析晶体结构的模块。它应用密度泛函理论，模拟计算各种材料（包括陶瓷、半导体、金属和生物系统）的体性质以及表面和界面的性质。它采用赝势平面波方法，密度泛函理论可以选用局

域密度近似或广义梯度近似，能够计算材料的系统总能量，原子间的相互作用力及材料的应力，对材料结构参数进行优化计算，进行分子动力学模拟研究，计算材料的弹性常数、晶格常数、能带结构、电子态密度、电荷密度分布、电子波函数以及光学性质。该程序可以在包括 Windows 和 Linux 等多种操作系统上运行，并且程序不断被更新。

CASTEP 实现运算的具体理论有：

1）离子−电子之间相互作用采用赝势来表示。

2）超晶胞的周期性边界条件。

3）平面波基组描述体系的电子波函数。

4）采用快速傅里叶变换（FFT）对体系哈密顿量进行数值化计算。

5）体系电子自洽能量最小化采用了迭代计算的方案。

6）采用交换−相关函数实现 DFT 的计算。

CASTEP 使用的是总能量平面波赝势方法，结合赝势和平面波基组的应用，使得对体系中所有原子上的力的计算极为容易。这就使得对分子固体表面以及界面的离子构型的有效优化成为可能。CASTEP 之所以功能如此强大的首要原因在于其求解电子方程的方法。

8.4.2 WIEN 软件包

WIEN 软件包是由维也纳技术大学（Vienna University of Technology）开发和编写的。它是一个采用全电子势线性缀加平面波方法加上局域轨道方法计算晶体性质的程序。 WIEN 程序所采用的全电子势考虑了电子的相对论效应，密度泛函理论可以选择局域密度近似、局域自旋密度近似和广义梯度近似。WIEN 可以计算晶体的能带结构和电子态密度、电子密度分布以及电子自旋密度分布、**X** 射线结构因子、费米面、晶体总能量、平衡态的结构参数，并可以通过分子动力学方法对结构参数进行优化，WIEN 还可以计算声子谱、**X** 射线发射和吸收谱、电子能量损失谱以及晶体的光学性质。WIEN 程序也可在多种操作系统下运行。

8.4.3 ABINIT 软件包

CASTEP 和 WIEN 都是付费软件，而 ABINIT 则是一个源程序完全公开的开放软件。它是由比利时的 Université Catholique de Louvain、Université de Liege、日本的 Mitsubishi Chemical Corp.、法国的 Commissariat Energie Atomique（Bruyeres）和美国的 Corning Inc.共同开发的一个软件包，ABINIT 针对不同的操作系统和计算机硬件环境分别编写了串行程序和并行程序，源程序可以从 ABINIT 的网站上直接下载，并提供有详细的使用说明和教程以及许多具体的计算实例。它在互联网上有一个论坛，所有 ABINIT 的用户和对它感兴趣的人可以在论坛中自由讨论

ABINIT 的使用技巧以及各种物理问题。ABINIT 利用赝势和平面波基矢组，在密度泛函理论框架内计算材料的总能量、电荷密度分布和电子结构。利用 ABINIT 进行固体电子结构计算时，有多种赝势可以选择，包括 Troullier-Martins（TM）赝势、Hartwigsen-Goedecker-Hutter（HGH）赝势、Goedecker-Teter-Hutter（GTH）赝势和 Fritz-Haber-Institute （FHI）赝势，对交换-关联能的近似也有局域密度近似、局域自旋密度近似和广义梯度近似的多种具体表达形式可供选择。ABINIT 软件包包括的功能有：根据应力进行结构参数优化计算，利用应力进行分子动力学模拟，构造动力学矩阵（dynamical matrices），布洛赫有效电荷和介电张量。除了主程序以外，软件包还提供一些应用工具程序：Mrgddb 和 Ifc 利用 ABINIT 的结果计算全声子带结构或热力学函数。Ifc 还用于计算原子力常数。Cut3d 用于对由 ABINIT 生成的三维密度或势文件进行后期处理。可以计算某一指定平面、轴线或点的电荷密度分布，也可以进行电荷密度的 Hirshfeld 分解。AIM 对由 ABINIT 生成的三维密度文件作后期处理，进行电荷密度的 Bader Atom-In-Molecule 分解。Conducti 用于计算频率依赖的光学传导率。Kptgen 用于寻找单位元胞中一组原子的对称性，生成电子 $\boldsymbol{k}$ 波矢格子，ABINIT 软件从开发初期到现在，源程序不断得到改进，功能不断得到加强。世界各地有许多研究小组都在使用 ABINIT 软件包进行研究工作，不少研究论文在 Nature、Science、Physical Review B 等杂志上发表。

8.5　第一性原理方法在材料科学研究中的应用

第一性原理允许科学家研究系统的电子、光学和结构性质的本质和根源，除了系统组成物质的原子序数外，并不需要任何实验数据。因此，非常适合解决固体物理、材料科学以及化学等领域中的问题。在这些领域的研究中，科学家可以应用计算机模拟进行虚拟实验，从而能大大节省实验费用并缩短研究周期。第一性原方法被广泛应用于表面化学、物理和化学吸附、多相催化、半导体缺陷、晶粒间界、堆垛层错、纳米技术、分子晶体、多晶型研究以及扩散机理等诸多领域。预测包括晶格参数、分子对称性、结构性质、能带结构、固态密度、电荷密度和波函数、光学性质等。下面列举一些实例，对第一性原理在材料科学研究中的应用作一简介。

8.5.1　Cu-Ni-Si 合金析出相稳定性的第一性原理研究

Cu-Ni-Si 系列合金的时效强化效应早在 1927 年就已经发现，但由于实验手段的限制，对其时效析出强化相的种类、结构一直存在着争论，几十年来，关于该类合金中析出物结构研究的文献连续出现，有报道称该析出物结构类似

$\gamma-Ni_5Si_2$，也有人认为其结构类似 $\beta-Ni_3Si$。下面利用第一性原理，分别计算 $\delta-Ni_2Si$、$\gamma-Ni_5Si_2$ 和 $\beta-Ni_3Si$ 的晶格常数、形成热、结合能以及态密度，分析它们的结构稳定性。

1. 计算模型与方法

$\delta-Ni_2Si$ 属于正交晶系，为 Co_2Si 型结构，空间群为 Pbnm，实验测定的晶格常数为 a＝0.372nm，b＝0.499nm，c＝0.706nm。其单胞中有 12 个原子，其中 8 个 Ni 原子和 4 个 Si 原子分别占据不同的 4e 位置，如图 8-6a 所示。$\gamma-Ni_5Si_2$ 属于三角晶系，为 $Ni_{31}Si_{12}$ 型结构，空间群为 P321，实验测定的晶格常数为 $a=b$=0.6671nm，c＝1.2288nm。其单胞中有 43 个原子，其中 Ni 原子 31 个，Si 原子 12 个，如图 8-6b 所示。$\beta-Ni_3Si$ 属于立方晶系，为 $AuCu_3$ 型结构，空间群为 Pm−3m，实验测定的晶格常数为 a＝0.3510nm。其单胞中有 4 个原子，其中 3 个 Ni 原子占据面心位置，1 个 Si 原子占据顶点位置，如图 8-6c 所示。

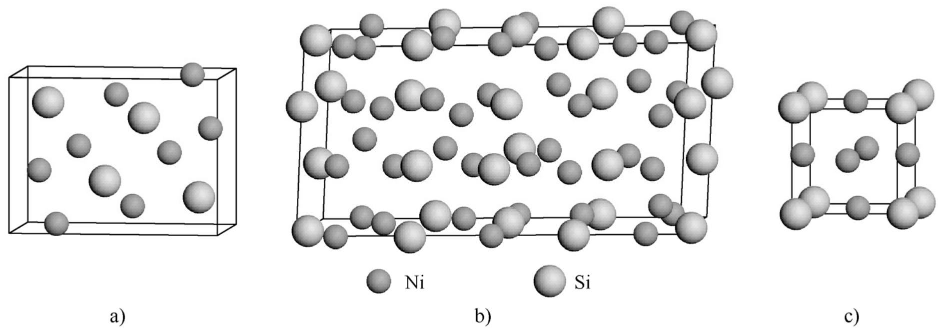


图 8-6 Cu-Ni-Si 合金中析出相的晶体结构模型

a）$\delta-Ni_2Si$ b）$\gamma-Ni_5Si_2$ c）$\beta-Ni_3Si$

采用基于密度泛函（DFT）理论的平面波赝势方法，计算程序是 CASTEP 软件包。晶体波函数由平面波基组来展开，交换关联能函数在结构优化中分别采用了 GGA 中的 PBE 和 LDA 中的 CA−PZ 形式，而其他计算只用 GGA 中的 PBE。采用周期性边界条件，平面波数目由动能截断点决定。所有结构计算的能量截断点取为 330 eV，$\delta-Ni_2Si$、$\gamma-Ni_5Si_2$ 和 $\beta-Ni_3Si$ 的 $\boldsymbol{K}$ 点设置分别为 $10\times8\times6$、$4\times8\times8$ 和 $6\times6\times6$，Ni、Si 元素都使用超软赝势，这些原子的赝势由 CASTEP 软件包提供。各项计算之前，先采用 BFGS 方法对超胞进行几何优化，以求得局域稳定结构。自洽场计算（SCF）时应用了 Pulay 密度混合法，SCF 的误差设置为 5×10^{-6} eV。自洽计算过程结束时，体系的总能量收敛到 5×10^{-6} eV/atom，每个原子上的力低于 0.1eV/nm 公差偏移小于 5×10^{-3}nm，应力偏差小于 0.02GPa。

2. 晶格常数

分别利用第一性原理的 GGA 和 LDA 方法对析出相 $\delta-Ni_2Si$、$\gamma-Ni_5Si_2$ 和

β−Ni_3Si 进行了结构优化，计算了各自的析出相晶格常数，其结果见表 8-1。

表 8-1　析出相晶格常数的第一性原理计算结果（GGA、LDA）

析出相	空间群/空间群数	GGA-PBE			LDA-CA-PZ			实验数据		
		a / nm	b / nm	c / nm	a / nm	b / nm	c / nm	a / nm	b / nm	c / nm
δ-Ni_2Si	Pbnm/62	0.7008	0.5123	0.3688	0.694	0.4891	0.3663	0.706	0.499	0.372
γ-Ni_5Si_2	P321/150	0.6614	0.6614	1.2233	0.6600	0.6600	1.2122	0.6671	0.6671	1.2288
β-Ni_3Si	$Pm\bar{3}m$/221	0.3512	—	—	0.3441	—	—	0.3510	—	—

从表 8-1 中可以看出，δ−Ni_2Si、γ−Ni_5Si_2 和 β−Ni_3Si 的计算结果都与实验结果非常接近，而且，所有结构的晶格常数的 GGA 计算结果都比 LDA 更接近于实验值。考虑到实验值一般是在 25℃下测试的，而第一性原理计算的条件为绝对零度（0K），因此，其计算结果是合理的、可靠的。

3. 形成热与结合能

形成热和结合能是固体稳定性判据的重要指标。形成热越大，热力学越稳定；结合能越大，晶体的结构越稳定。晶体的结合强度和结构稳定性与其结合能密切相关，晶体结合能就是自由态原子结合成晶体的能量，或者是晶体分解成单个自由态原子所做的功。A_mB_n 二元合金相的单位原子结合能 $E_{coh}^{\varphi}(A_mB_n)$ 可以用下式来表示

$$E_{coh}^{\varphi}(A_mB_n) = E^{\varphi}(A_mB_n) - \frac{1}{m+n}\left[mE_{atom}^{A} + nE_{atom}^{B}\right] \tag{8-58}$$

式中，$E^{\varphi}(A_mB_n)$ 表示结构为 φ 的二元合金相单胞或原胞中每个原子的能量，E_{atom}^{A} 和 E_{atom}^{B} 分别表示单个自由态 A 和 B 原子的能量。

A_mB_n 二元合金相的平衡态单位原子形成热 $\Delta H^{\varphi}(A_mB_n)$ 可以用下式来表示

$$\Delta H^{\varphi}(A_mB_n) = E^{\varphi}(A_mB_n) - \frac{1}{m+n}\left[mE^{\alpha}(A) + nE^{\beta}(B)\right] \tag{8-59}$$

式中，$E^{\alpha}(A)$表示为 α 结构的 A 原胞中每个原子的能量；$E^{\beta}(B)$表示为 β 结构的 Ni 原胞中每个原子的能量。

利用第一性原理计算了 δ−Ni_2Si、γ−Ni_5Si_2 和 β−Ni_3Si 的单个原子平均单胞总能量（E_T)、形成热（ΔH）和结合能（E_{coh}），其计算结果如表 8-2 所示。

表 8-2　析出相的单个原子平均单胞总能量、形成热和结合能的计算结果

析出相	单个原子总能量 E^T、形成热 ΔH 和结合能 E_{COH} 的第一性原理计算结果（GGA-PBE）		
	E^T/(eV/atom)	ΔH/(kJ/mol)	E_{COH}/(eV/atom)
δ−Ni_2Si	−941.125	−54.909	−6.876
γ−Ni_5Si_2	−1008.776	−48.416	−6.871
β−Ni_3Si	−1045.009	−45.204	−6.870

从表 8-2 中可以看出，δ−Ni_2Si 形成热最大，为−54.909 kJ/mol；γ−Ni_5Si_2 次之，为−48.416 kJ/mol；β−Ni_3Si 最小，为−46.362 kJ/mol。三种析出相的形成热都很大，从热力学观点出发，它们都是非常稳定的相结构。通过比较可得，δ−Ni_2Si 最稳定，其次为 γ−Ni_5Si_2，最后为 β−Ni_3Si。另外，从表中还可以看出，三种结构的单个原子结合能计算结果十分接近（−6.876eV/atom、−6.871eV/atom、−6.870eV/atom），可以认为是相等的，三种结构析出相的结构稳定性没有明显差别，δ−Ni_2Si 的略好。

综合考虑形成热和结合能两种因素，三种析出相在热力学上都是非常稳定的结构，从高到低的结构稳定次序为：δ−Ni_2Si＞γ−Ni_5Si_2＞β−Ni_3Si，与有关的实验和理论计算结果基本一致。

4. 态密度

态密度是一个十分有用的概念，使用态密度可以用对电子能量 E 的积分代替在对布里渊区内对 $\boldsymbol{k}$ 的积分。另外，态密度经常用于电子结构的快速可视分析，价带宽度、能隙及电子态密度 $N(E)$ 的主要特征处的强度和数目等特性有助于定性解释实验得到的光谱数据。态密度分析还有助于理解电子结构的变化，如外压引起的电子结构变化。

局域态密度（LDOS）和分波态密度（PDOS）是对电子结构分析十分有用的半定量工具。LDOS 显示系统中各原子的电子态对能态密度谱的每个部分的贡献。PDOS 根据电子态的角动量来进一步分辨这些贡献，确定 DOS 的主要峰是否具有 s、p 或 d 电子的特征。LDOS 和 PDOS 分析可对体系中电子杂化的本质和体系的 XPS 谱、光谱中主要特征的来源提供定性解释。

δ−Ni_2Si、γ−Ni_5Si_2 和 β−Ni_3Si 的态密度计算结果分别如图 8-7～图 8-9 所示。图 8-7～图 8-9a、b 和 c 分别为 δ−Ni_2Si、γ−Ni_5Si_2 和 β−Ni_3Si 的总态密度（TDOS）以及 Ni 和 Si 原子的分波态密度（PDOS），图中虚线代表费米面的位置。为了清楚显示 Ni 原子的 s 和 p 态的 PDOS，图 a 中 Ni 的 d 态 PDOS 的峰值被截断。实际上，结构的成键态主要来自 Ni 的 d 态的峰值贡献，在费米面附近，PDOS 主要由 p 态和 d 态组成，s 态的作用可以忽略不计。

从整体上看，三种析出相的 DOS 曲线基本相似。一个明显的特征就是 Ni−3d 态与 Si−3p 态电子的强烈杂化，δ−Ni_2Si、γ−Ni_5Si_2 和 β−Ni_3Si 都有一个尖锐的强键峰，分别位于−5.4～−0.3eV、−5.3～0.5 eV 和−4.1～0.1 eV 区域内。另外一个特征就是在费米面附近有一个不明显的赝能隙（pseudogap），把 p 态和 d 态的成键态和反键态分开。

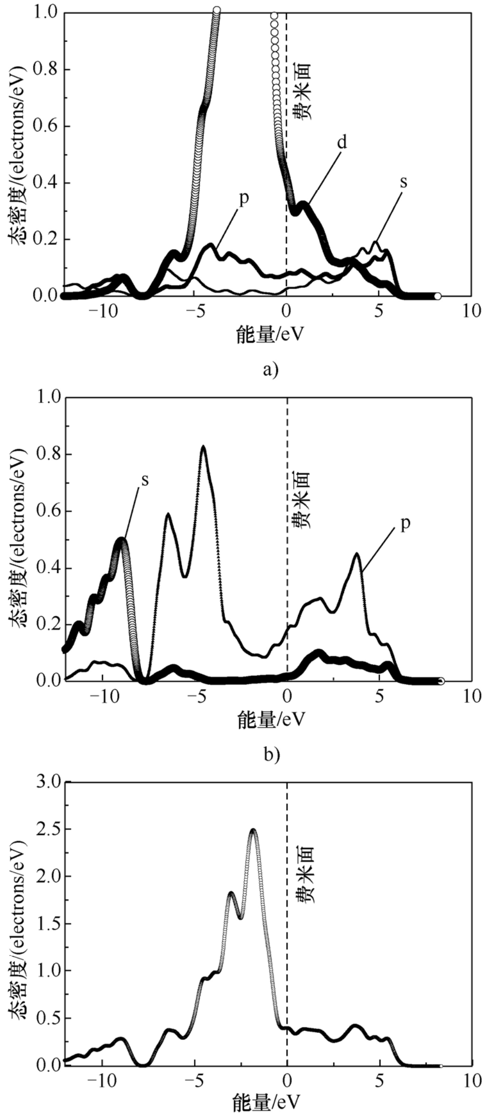


图 8-7　δ−Ni_2Si 的总态密度以及其 Ni、Si 原子的分波态密度

a）Ni 的 PDOS　b）Si 的 PDOS　c）δ-Ni_2Si 的总态密度

通过仔细比较分析三种结构的 TDOS 图，可以得出几点：

1）费米面以下 δ−Ni_2Si 的强主峰位置（−1.75～−2.91eV）比 γ−Ni_5Si_2（−1.16～−2.71eV）和 β−Ni_3Si（−1.12～−2.90eV）向左有所移动，说明 δ−Ni_2Si 结构更稳定。

2）δ−Ni_2Si 的费米面处的 DOS 较低（0.41states/eV），费米面几乎位于赝能隙

的谷底，而 γ−Ni_5Si_2 和 β−Ni_3Si 的赝能隙位于强主峰的右坡面。赝能隙位于谷底的结构更稳定。

3）δ−Ni_2Si 的赝能隙宽度与其他两种结构相比也较大，赝能隙越宽，则系统共价成键性越强，结构也越稳定。

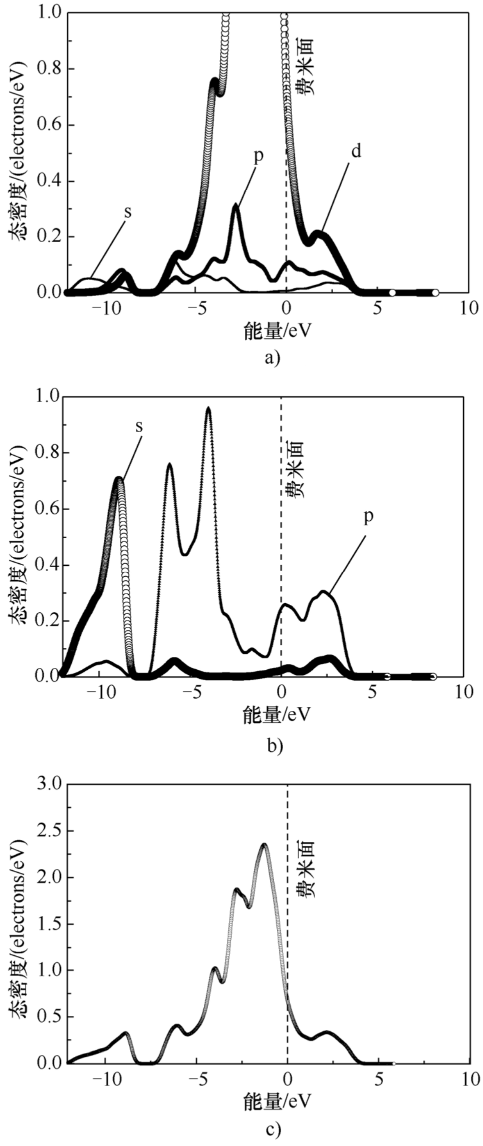


图 8-8　γ−Ni_5Si_2 的总态密度以及其 Ni 和 Si 原子的分波态密度

a）Ni 的 PDOS　b）Si 的 PDOS　c）γ−Ni_5Si_2 的总态密度

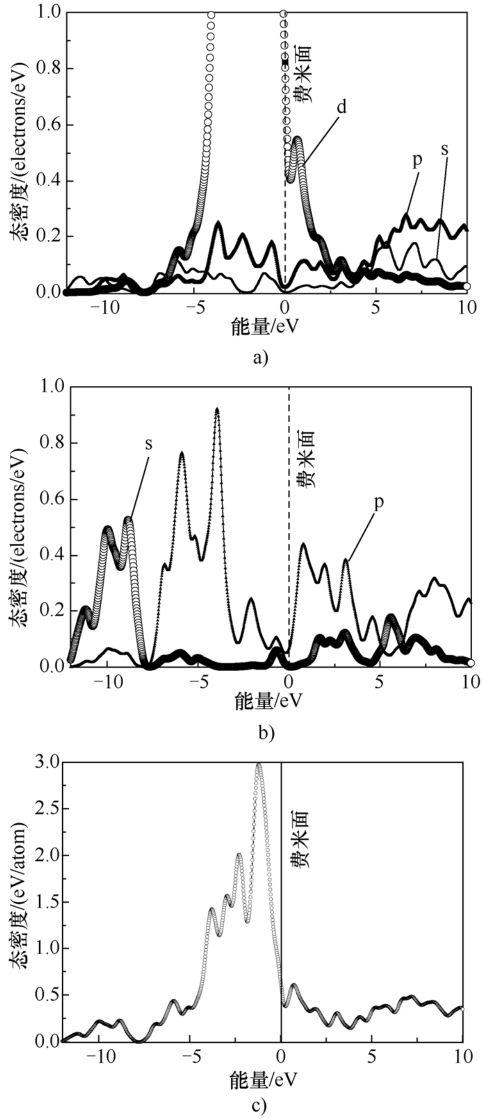


图 8-9　β−Ni_3Si 的总态密度以及其 Ni、Si 原子的分波态密度

a）Ni 的 PDOS　b）Si 的 PDOS　c）β−Ni_3Si 的总态密度

8.5.2　BN 电子结构的第一性原理计算

1. BN 的能带结构

全部 $E_n(\boldsymbol{k})$函数给出能带结构，实际上人们最关心的是费米能级附近的一系列能带，因此，一般只计算有限个能带的 $E_n(\boldsymbol{k})$函数。从能带结构图上可以直观地看

到在指定方向上各能带 $E_n(\boldsymbol{k})$函数随 $\boldsymbol{k}$ 的变化、导带底与价带顶的位置、禁带宽度以及禁带能隙随 $\boldsymbol{k}$ 的变化。

图 8-10 是利用 CASTEP 软件计算的闪锌矿结构的 BN 的能带结构图，总能量的计算精度是1.0×10^{-5}，$\boldsymbol{k}$ 的取值为 8×8×8，能量截断值为 250.0 eV。图中用虚线表示费米能级（$E_F = 0$），费米能级以下的能带是价带，费米能级以上的能带是导带。导带底与价带顶位于相同点的能带结构，称为直接跃迁型能带结构。导带底与价带顶位于不同点的能带结构，称为间接跃迁型能带结构。从图 8-10 中可以看出，导带底与价带顶均位于布里渊区的原点（Γ 点，图中以 G 表示 Γ），这表明 BN 具有直接跃迁型能带结构。

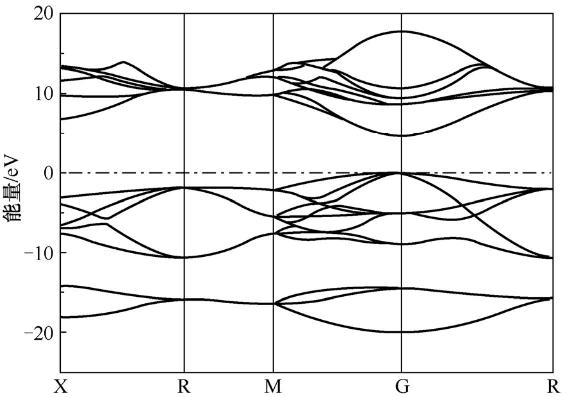


图 8-10　BN 的能带结构

2. 态密度

BN 的态密度计算结果如图 8-11～图 8-13 所示。图 8-11～图 8-13 分别为 BN 的总态密度、BN 中 B 原子和 N 原子的局域态密度、BN 中 B 原子和 N 原子的分波态密度。

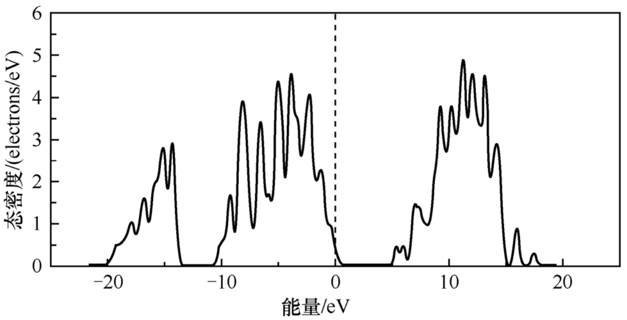


图 8-11　BN 的总的能态密度

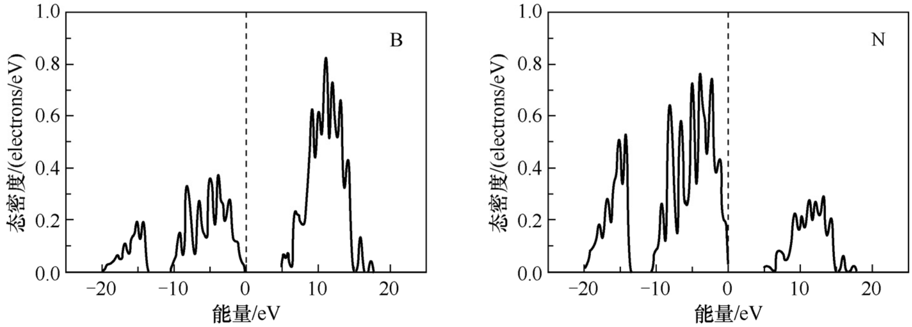


图 8-12 BN 中 B 原子和 N 原子的局域态密度

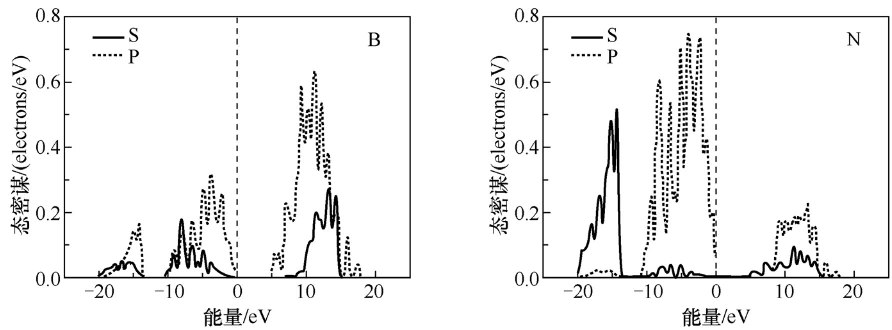


图 8-13 BN 中 B 原子和 N 原子的分波态密度

3. 布居分析

对电子电荷在各组分原子之间的分布情况进行计算，称为布居分析。由布居分析得到的原子电荷值只有相对意义没有绝对意义，因为原子电荷值对所用基组十分敏感。如果使用相同的基组对不同体系进行布居分析计算，得到的电荷分布的相对值可以给出一些有用的信息。有多种布居分析方法，其中被广泛采用的布居分析方法是 Mulliken 布居分析。布居分析可以给出原子上、原子轨道上、两原子间的电子电荷分布，依次称为原子布居、轨道布居、键布居。

布居分析为原子间的成键提供了一个客观判据，并且两原子间的重叠布居还可用于评价一个键的共价性或离子性。键布居的值高表明键是共价的，键布居的值低表示键是一种离子相互作用。还可以用有效离子价来进一步评价键的离子性，有效离子价定义为阴离子物种上原来的离子电荷与 Mulliken 电荷之差，若这个值为零，则表明该键是完全的离子键，若这个值大于零，则表明该键的共价成分增加。

对 BN 做 Mulliken 布居分析，得到 B 和 N 的原子布居列于表 8-3 中，表中列出了 B 原子和 N 原子的轨道布居、总的原子布居及原子电荷值。

表 8-3 BN 的 Mulliken 布居分析结果

原子	s 轨道	p 轨道	总的原子布居	原子电荷值（e）
B	0.65	1.75	2.40	0.60
N	1.50	4.10	5.60	−0.60

8.5.3 TiB 弹性性质的第一性原理计算

TiB 因其具有高的弹性模量而在钛基复合材料中作为增强体，备受重视。一般材料的弹性模量可以通过实验方法测得，但对于难以测试的块体材料（如 TiB）或材料中析出相的弹性模量存在一定困难。近几年来，由于第一性原理对材料弹性性质的理论计算结果与实验值有很好的一致性，因此已成为研究材料弹性性质的重要手段。

TiB 属于正交晶系，其空间群为 P6/mmm，单胞中含有 3 个原子，具有 9 个独立的弹性常数（C_{11}，C_{22}，C_{33}，C_{12}，C_{13}，C_{23}，C_{44}，C_{55} 和 C_{66}）。下面是利用基于密度函数理论和赝势平面波方法计算得到的 TiB 弹性性质。

采用计算得到的弹性常数可以计算出 TiB 的体模量 B、弹性模量 E、切变模量 G 和泊松比 v 等弹性性质参数。在多晶材料弹性模量的理论估算方面，Hill 通过极值原理证明，Voigt 和 Reuss 模型的计算结果是弹性常数的上下限。Hill 模型则将 Voigt 和 Reuss 模型的计算结果取一个简单的算术平均即 VRH 平均，其结果和实际测定值更为一致。Hill 模型即 VRH 平均法计算多晶体材料体模量 B 和切变模量 G 的计算式为

$$B_H = \frac{1}{2}\left(B_R + B_V\right) \tag{8-60}$$

$$G_H = \frac{1}{2}\left(G_R + G_V\right) \tag{8-61}$$

采用 Voigt 和 Reuss 模型计算正交晶系晶体的 K 和 G 的计算为

$$B_V = \frac{1}{9}\left(C_{11} + C_{22} + C_{33}\right) + \frac{1}{9}\left(C_{12} + C_{23} + C_{13}\right) \tag{8-62}$$

$$B_R = \frac{1}{\left(S_{11} + S_{22} + S_{33}\right) + \left(S_{12} + S_{23} + S_{13}\right)} \tag{8-63}$$

$$G_V = \frac{1}{15}\left(C_{11} + C_{22} + C_{33}\right) - \frac{1}{15}\left(C_{12} + C_{23} + C_{13}\right) + \frac{1}{5}\left(C_{44} + C_{55} + C_{66}\right) \tag{8-64}$$

$$G_R = \frac{15}{4\left(S_{11} + S_{22} + S_{33}\right) - 4\left(S_{12} + S_{23} + S_{13}\right) + 3\left(S_{44} + S_{55} + S_{66}\right)} \tag{8-65}$$

式中，S_{ij} 为 C_{ij} 的逆矩阵。弹性模量 E 和泊松比 v 的计算式为

$$E = \frac{9BG}{\left(3B + G\right)} \tag{8-66}$$

$$v = \frac{(3B - E)}{6B} \tag{8-67}$$

采用上述公式用表 8-4TiB 材料的弹性常数可计算出 TiB 材料的弹性模量 *E*、切变模量 *G*、体模量 *B* 和泊松比 *v*（见表 8-5）。

表 8-4　TiB 晶体的弹性常数（GPa）

C_{11}	C_{22}	C_{33}	C_{12}	C_{13}	C_{23}	C_{44}	C_{55}	C_{66}
425	529	409	86	106	61	201	182	224

Pugh 曾提出过一个预测材料的延/脆性的经验判据，即通过切变模量 *G* 与体模量 *B* 的比值来判断金属材料的延性或脆性。如果这一比值 $G/B < 0.5$，材料呈延性；反之，材料则呈脆性。这一判据已被广泛应用于分析金属间化合物和类金属间化合物的延性或脆性。从表 8-5 可以看出，TiB *G/B* 值远大于 0.5，说明 TiB 是具有很大脆性的化合物。另外，泊松比 *v* 的大小也可用来衡量金属和金属间化合物的脆性。对延性材料而言，泊松比一般为 1/3；而脆性材料的泊松比一般小于 1/3。由表 8-5 可知，材料 TiB 的泊松比远小于 1/3，所以再次说明 TiB 是具有很大脆性的化合物材料。

表 8-5　TiB 材料的体模量 *B*（GPa）、切变模量 *G*（GPa）、弹性模量 *E*（GPa）和泊松比 *v*

B_V	B_R	B_H	G_V	G_R	G_H	E	G/B	v
207	206	206.5	186	183	184.5	427	0.89	0.15

第 9 章　原子间相互作用势

原子间相互作用势是所有有关原子水平上的计算机模拟的基础，其势的精确与否将直接影响着模拟结果的准确性，而计算机模拟所需要的计算机机时则取决于势函数的复杂程度。如果从第一原理出发，对某一材料进行完全的量子力学处理，不仅在计算方法上存在一定的困难，而且难以获得全面而准确的计算结果。在一定的物理模型的基础上发展相应的原子间相互作用势，进而研究材料的性质和不同状态下的行为，成为材料研究中一种必要的研究手段。早期的原子间相互作用势多数是一些纯经验拟合势，近年来人们更多的是通过基本电子结构的理论计算，发展一些合适的半经验的“有效势”。

本章对材料计算中广泛使用的原子间相互作用势方法进行了概述，其中重点介绍了嵌入原子势方法。

9.1　原子间相互作用势概述

就描述对象的物理实质深刻程度来说，原子模拟计算介乎宏观的热力学理论和量子力学第一原理理论之间。由于原子模拟较好地考虑了原子间的相互作用，加之不涉及到解薛定谔方程，因而相对比较简单，它能很好地描述许多固体，特别是金属和合金问题，得到了与实验结果符合得好的结果。因此原子模拟理论得到了重视和发展，在材料设计中成了不可或缺的理论。但是由于实验上无法直接测定原子间的相互作用，因此对原子间势的表述就是关键。人们选取的原子间势不同，原子模型也就不同。如果模型太简单，那么就难于较精确地描述实际的金属及合金；如果太复杂，又难于较广泛地进行有效的应用。

原子间相互作用势是原子级材料计算和设计的基础。从 20 世纪 20 年代开始，人们就提出了各种各样的描述原子间相互作用的函数形式，如早期著名的 Lennard–Jones 势、Morse 势、Buckingham 势、Bom–Mayer 势等原子间相互作用对势形式和近 20 年发展起来的 Tersoff、Stilhnger–Weber、Embedded–atom method（EAM）等原子间相互作用多体势形式，它们在材料的缺陷、表面和界面、断裂、团簇的结构及其它们在温度特性的计算模拟中起着重要的作用。

随着量子力学理论和计算技术的发展以及计算机性能的飞速提高，对材料结构和特性量子力学的研究广泛地开展起来。但由于计算条件的限制，对于由 10^3 个以上的原子组成的复杂体系的各种特性的研究，目前仍然难以通过第一原理计

算完成。因此，在材料的静力学、动力学以及统计力学的计算中，各种各样等效的原子间相互作用势仍然起着重要的作用。在对材料从微观到宏观的某些过渡尺度对象的研究中，原子间相互作用势仍然具有不可替代的作用。近二十年来，基于第一原理计算的原子间相互作用势，由于其强大的理论背景而日益受到人们的重视。人们在不断地探索如何基于第一性原理计算，准确而又简洁地获得原子间相互作用势的方法。

20 世纪 90 年代，陈难先先生创造性地将数论中的 Möbius 变换应用到凝聚态物理的研究，尤其是在三维晶格反演中的应用，给出了一条从第一原理结合能曲线得到原子间相互作用对势的解析途径，由此得到的原子间相互作用对势广泛地应用到了金属、合金、金属间化合物和稀土过渡金属磁性材料的结构、静力学和动力学特性的研究。

9.2　原子间相互作用对势的形式

在材料的静力学、动力学以及统计力学的计算中，普遍采用的原子间相互作用势经历了很长的发展过程，从最近邻的原子对相互作用到复杂的多体相互作用势，它们分别在不同的领域，对不同的物理对象的处理发挥着重要的作用。下面分别介绍一些常用的原子间相互作用对势的形式。它们常用于处理一些对原子间角度不敏感的、或高对称性的、或稀薄气体等物理对象。

9.2.1　“刚球”模型

最简单的原子模型是一个“刚球”模型，函数形式为

$$V(r)=\begin{cases}\infty, & r<\sigma \\ 0, & r>\sigma\end{cases} \tag{9-1}$$

式中，参数 σ 是原子半径。

这种势形式的最大特点是简单，它作为原子间相互作用势一级近似广泛地应用于一些唯象的研究。在某些情况下，需要原子间相互作用势的参与，但对它的精度并没有较高的要求时，例如，研究固体中由于高能辐射引起的原子级联错位，这时，刚球模型可以对每个高能粒子所引起的错位原子的能量和数量分布给出一个数量级的估算。相对于具体的碰撞事例的能量，可以通过改变刚球半径；来改善势参数的有效性。

刚球模型的修正形式是通过方势阱引入原子间相互作用的吸引部分

$$V(r)=\begin{cases}\infty, & r<\sigma \\ -\varepsilon, & \sigma<r<R\sigma \\ 0, & r>R\sigma\end{cases} \tag{9-2}$$

这是一个刚球模型上围上了一层深度为 ε、厚度为 $\sigma(R-1)$的吸引壳层。由于它增加了两个可调参数 ε 和 R，因此可以通过一定的调整，使之更加符合实验数据。

取代方势阱，可以考虑原子不是“全硬”的，原子之间存在着如下的原子间相互作用

$$V(r)=\frac{\lambda}{r^n} \tag{9-3}$$

式中，n 的取值在 9～15 之间。这种势被称为玻恩（Born）势，它是许多其他复杂经验势函数的基础。这里，原子是斥力中心，根据 n 和 λ 的值能够给出依赖于相互作用能量的原子有效半径。这种势常常用于计算原子间的相互作用力。

刚球模型的另一个修正是计入如下的原子间相互吸引

$$V(r)=\begin{cases}\infty, & r<\sigma \\ -\dfrac{\lambda}{r^n}, & r>\sigma\end{cases} \tag{9-4}$$

式中，n 的取值为 5 或 6。这种势被称为 Sutherland 势，它保持了数学形式上的简洁性，更加逼近真实的物理模型

以上介绍的势形式仅仅计及中性原子之间的相互作用。荷电粒子之间的相互作用需要加入 Coulomb 相互作用的影响。

9.2.2 Buckingham 势

Buckingham 提出了包含偶极和四极相互作用的势形式，他将斥力展开成指数项

$$V(r)=A\mathrm{e}^{-Br}-\frac{\lambda}{r^6}-\frac{\lambda'}{r^8} \tag{9-5}$$

从理论的观点考虑，与“硬球”模型相比，这种势更加接近于真实势，但数学的处理上较为复杂。另外，它有一个明显的缺陷，就是当两个原子的距离趋于零时，势趋于负无穷大。

Buckingham−Corner 势克服了这种不真实的状况，势形式被修正为如下的复杂形式

$$V(r)=\begin{cases}A\exp\left\{-\alpha\left(\dfrac{r}{r_m}\right)\right\}-\left(\dfrac{\lambda}{r^6}+\dfrac{\lambda'}{r^8}\right)\exp\left\{-4\left(\dfrac{r_m}{r}-1\right)^3\right\}, & r<r_m \\ A\exp\left\{-\alpha\left(\dfrac{r}{r_m}\right)\right\}-\left(\dfrac{\lambda}{r^6}+\dfrac{\lambda'}{r^8}\right), & r\geqslant r_m\end{cases} \tag{9-6}$$

式中， $A=\left\{-\varepsilon+\left(1+\beta\right)\left(\lambda/r_m^6\right)\right\}\mathrm{e}^{\alpha}$ ；

$\lambda=\varepsilon\alpha r_m^6\Big/\left\{\alpha\left(1+\beta\right)-6-8\beta\right\}$ ；

$\lambda'=\beta\lambda r_m^2$ 。

ε 是势能的极小值，r_m 是与之相对应的 r 值。α 反映了指数项的“陡峭”程度，β 是 $r = r_m$ 时 r 的 8 次方项与 6 次方项贡献的比值。它有 4 个可调参数可用于拟合实验数据。从解析的角度看，这种势的形式过于复杂，难以调整。

一个相关的简单势形式是修正的 Buckingham 势，它削弱了 8 次方项的影响

$$V(r)=\frac{\varepsilon}{1-6/\alpha}\left\{\frac{6}{\alpha}\exp\left[\alpha\left(1-\frac{r}{r_m}\right)\right]-\left(\frac{r_m}{r}\right)^6\right\} \tag{9-7}$$

式中，各符号类似于式（9-6）。这种势在 $r_{\max}$ 有一个最大值，$r_{\max}$ 由下式决定

$$\left(\frac{r_{\max}}{r}\right)^7\exp\left\{\alpha\left(1-\frac{r_{\max}}{r}\right)\right\}=1 \tag{9-8}$$

这显然不符合相互作用的实际情况，但可以通过一个简单的假设解决这一问题，即：$V(r)=\infty$，$r\leqslant r_{\max}$。由于 $r_{\max}$ 通常足够得小，导致这一假设是可以接受的。偶极矩感应诱导的四极矩相互作用项（8 次方项）可以通过 α 的微小变化反映出来。

9.2.3　Lennard–Jones（L–J）势

Mie 最先提出了两体式的解析形式，Lennard–Jones 将它应用到铜的自扩散的计算中。Lennard–Jones 势的早期形式为 $V(r)=4\varepsilon[(\sigma/r)^{12}-(\sigma/r)^6]$，由于其排斥项和吸引项的幂分别为 12 和 6，因此又称之为“6-12”次方势。L–J 势对惰性气体元素原子间的相互作用的描述比较好，也可以用来描述原子中电子在空间中的分布方向性不太强的元素，如铜、银、金等。在描述这些元素的原子间相互作用时，通常为更好地拟合元素已有的实验数据，如结合能、晶格常数等，Lennard–Jones 势则使用一种普适的一般形式

$$V(r)=4\varepsilon\left(\frac{n}{m-n}\right)\left[\left(\frac{\sigma}{r}\right)^m-\left(\frac{\sigma}{r}\right)^n\right] \tag{9-9}$$

式中，m、n、ε 和 σ 为势参数，r 表示原子间距。

多年来，Lennard–Jones 势依然是人们定性地研究材料的各种结构、力学特性、动力学特性和热力学特性时，广泛采用的原子间相互作用势模型。

9.2.4　Morse 势

为了计算分子能级和光谱，Morse 采用了完全指数形式的函数来描述双原子分子的原子之间的相互作用

$$V(r)=D_e\mathrm{e}^{-2\alpha(r-r_e)}-2D_e\mathrm{e}^{-\alpha(r-r_e)} \tag{9-10}$$

式中，r 是两原子间距或键长，r_e 为两原子平衡间距或平衡键长（$\mathrm{d}V(r)/\mathrm{d}r\big|_{r=r_e}=0$），$D_e$ 和 α 是模型参数（D_e 称为 Morse 的势阱深，α 则控制着势阱的“宽度”，α 越小，则势越宽。）

Morse 势并不只限于用来解释分子的能级，Girifalco 和 Weizer 最早用它计算了立方晶系的原子间相互作用势参数。根据 Morse 势函数，晶体的每原子的平均结合能

$$E=\frac{D_e}{2}\sum_{ij}\left\{e^{-2\alpha(r_{ij}-r_e)}-e^{-\alpha(r_{ij}-r_e)}\right\} \tag{9-11}$$

设系统的压缩比为 K_B，则 $\frac{1}{K_B}=\left[V\bullet\frac{d^2E}{dr^2}\right]_a$。由此，可以根据晶格常数、结合能、压缩比（或体模量）确定 Morse 势参数 D_e、α 和 r_e。

Girifalco 和 Weizer 通过计算给出了 16 种纯金属的 Morse 势参数，并使用 Debye 模型处理自由能中的热学部分，结合 Morse 势参数，得到的这些金属的状态方程曲线与实验结果一致。另外，他们计算了弹性常数在有些事例中，与实验值较为一致。

迄今为止，Morse 势函数依然广泛地应用于描述金属或合金中的原子间相互作用。

9.2.5 Bom-Mayer 势

Bonm 和 Mayer 提出了一个简单的两参数指数形式的势

$$V(r)=Ae^{-Br} \tag{9-12}$$

式中，A 和 B 为模型参数。

Bonm 和 Mayer 利用上式，处理了离子晶体中闭壳层的离子之间的排斥作用。尽管它形式简单，但它是一个基于纯理论的考虑，这个指数形式得到量子力学理论的支持。

当 r 趋于零时，从势的有限性可以看出，它的有效性仅仅局限于闭壳层的电子云中等程度交叠的情况。实际上，常常通过平衡情况下的晶格数据确定势参数 A 和 B，因此严格地讲，它的正确性仅仅局限于闭壳层的电子云轻度交叠的情况。

Huntington 对上述的势形式稍加改进，并将其应用到晶格系统。它将平衡的最近邻距离 r_0，引入到势函数中

$$V(r)=A'\exp\left\{-\rho(r-r_0)/r_0\right\} \tag{9-13}$$

这种势函数的形势，目前仍然广泛地应用于描述离子晶体中，离子之间短程排斥相互作用。

9.2.6 多项式形式的势

Johnson 在处理 Fe 和 V 的相互作用时，采用了如下形式的原子间相互作用势

$$V(r)=A(r-B)^3-Cr+D \tag{9-14}$$

并将其应用到 α-Fe 的空位附近原子组态的计算机模拟之中。

类似于式（9-14）形式的原子间相互作用势，由于其可调参数较多，人们多用来拟合较为复杂的物理现象，但也由于其参数较多，失去了某些物理意义上的简洁性。

以上介绍了一些典型的原子间相互作用势的形式。基于当时的理论和计算的条件，最初这些势的参数的确定大多是根据实验数据来确定。通常称之为经验的原子间相互作用势。

9.3　原子间相互作用多体势模型

一般而言，原子间的相互作用与环境有关，亦即，原子间的相互作用势应该包含多体关联效应，尤其是在处理凝聚态材料时。对势模型由于其简洁性在近一个世纪的研究中，广泛地被物理学家、化学家和材料科学家所关注，但它并不能反映原子间相互作用的多体效应的细节，应用范围受到了很大的局限。同时，晶体结构的稳定性和立方晶系弹性常数的 Cauchy 关系也是它无法克服的问题。

为了克服原子间相互作用对势这些局限性，各种经验的、半经验的以及第一原理模式的原子间相互作用多体势模型被提出，在这些模型中，引入了依赖于体积或电子密度的非中心力部分，用于描述原子间相互作用的角度因素，诸多反映多体效应的形式，诸如 Stillinger-Weber 多体势、Daw 和 Baskes 的埋入原子势（EAM、Finnis-Sinclair（F−S）多体势、Tersoff−Brenner 的多体势、Finnis−Sinclair 多体势、Slater−Kirkwood 类型的三体势等，已成功地应用于一些材料问题的研究。

除了嵌入原子势（EAM）和 Finnis−Sinclair（F−S）多体势（内容太多，另作介绍）以外，其他的多体势介绍如下。

9.3.1　Tersoff 和 Brenner 势

Tersoff 势和 Brenner 势的一般形式为

$$E_{\text{total}} = \frac{1}{2}\sum_{i\neq j} V_{ij} = \frac{1}{2}\sum_{i\neq j}[V_R(r_{ij}) + B_{ij}V_A(r_{ij})] \tag{9-15}$$

其中，V_R 是排斥项，V_A 是吸引项，B_{ij} 是一个与键角 θ_{ijk} 有关的系数。

1. Tersoff 势

J. Tersoff 在 1986 年提出了一个描述共价晶体原子之间相互作用的势具体形式

$$V_{ij}(r) = f_c(r_{ij})\left[A\exp(\lambda_1 r_{ij}) - B_{ij}\exp(\lambda_2 r_{ij})\right] \tag{9-16}$$

式中，$V_{ij}(r)$是 i 和 j 原子之间的相互作用能，r_{ij} 是原子间距，A、B_{ij}、λ_1 和 λ_2 是正的参数，且有 $\lambda_1 > \lambda_2$。f_c 是一个任意截断函数，它限制了势的作用范围。式中第一项是排斥相互作用，它的起源通常解释为由于波函数的正交化等，第二项反映化学键的效应。B_{ij} 反映键序（bond order）和周围环境的影响

$$B_{ij} = B_0 \exp(-z_{ij}/b)$$
$$z_{ij} = \sum_{k \neq i,j} \left[w(r_{ik})/w(r_{ij}) \right]^n \left[c + \exp(-d \cos\theta_{ijk}) \right]^{-1} \tag{9-17}$$

这里，$w(r)$是“裸”的键势，$w(r) = f(r)\exp(-\lambda_2 r)$。反映的是 ik 键的数量对 ij 键的影响，它是 ik 键的数量与键 ij“竞争”的加权度量，b 是反映键强度随有效作用距离衰减的因子。式中第一项反映的是 ik 和 ij 键强度的无量纲的比值。θ_{ijk} 是 ij 和 ik 键之间的夹角。在他们的工作中

$$f_c = \begin{cases} 1, & r < (R+D) \\ \dfrac{1}{2} - \dfrac{1}{2}\sin\left[\dfrac{1}{2}\pi(r-R)/D\right], & (R-D) < r < (R+D) \\ 0, & r > (R+D) \end{cases} \tag{9-18}$$
$$R = 3.0, \quad D = 0.2$$

它们结合晶格常数、结合能和体模 f 给出了 Si 的势参数，并计算了若干不同结构的键能和结构几何，包括表面能和缺陷结构等。另外，它们将这种势形式应用到碳，计算了它的结构、能量、弹性、声子谱等，并把势应用于无定形碳，得出了类似于 α−C 的结构。

这种势在金刚石结构 Si 的稳定性方面并不成功，它给出 BCC 是更加稳定的结构。在 Tersoff 后续的工作中，势的函数形式修正如下

$$V_{ij} = f_C(r_{ij})\left[a_{ij} f_R(r_{ij}) + b_{ij} f_A(r_{ij}) \right]$$
$$f_R(r) = A\exp(-\lambda_1 r)$$
$$f_A(r) = -B\exp(-\lambda_2 r)$$
$$f_C = \begin{cases} 1, & r < (R+D) \\ \dfrac{1}{2} - 1\sin\left[\dfrac{1}{2}\pi(r-R)/D\right], & (R-D) < r < (R+D) \\ 0, & r > (R+D) \end{cases}$$
$$b_{ij} = (1 + \beta^n \zeta_{ij}^n) \tag{9-19}$$
$$\zeta_{ij}^n = \sum_{k(\neq i,j)} f_C(r_{ik}) g(\theta_{ijk}) \exp\left[\lambda_3^3 (r_{ij} - r_{ik})^3 \right]$$
$$g(\theta) = 1 + c^2/d^2 - c^2/\left[d^2 + (h - \cos\theta)^2 \right]$$
$$a_{ij} = \left(1 + \alpha^n \eta_{ij}^n\right)^{-1/2n}$$
$$\eta_{ij}^n = \sum_{k(\neq i,j)} f_C(r_{ik}) \exp\left[\lambda_3^3 (r_{ij} - r_{ik})^3 \right]$$

式中，i、j 和 k 标记系统中的原子，其余标号的意义与式（9-16）相同。改进后的原子间相互作用势能够给出稳定的金刚石结构的 Si，能够很好地描述它的声子谱和弹性，对液态 Si 的描述也得到了改进，但对表面的描述并不精确。这种形式的

势也应用于研究金刚石 C 的各种特性及其无定型 C。

另外，Tersoff 势的各种修正形式也广泛地应用于金刚石结构 Si 的研究，如 Dodson 势

$$
\begin{aligned}
&U_{ijk} = f_{ij}(A\mathrm{e}^{-2\lambda r_{ij}} - B_{ij}\mathrm{e}^{-2\lambda r_{ij}}) \\
&B_{ij} = B_0\mathrm{e}^{-z_{ij}^{n}/b} \\
&z_{ij} = \sum_{k\neq i,j}\left[\frac{f_{ij}}{f_{ik}}\right]^4 \mathrm{e}^{4\lambda(r_{ij}-r_{ik})}\left[c+\mathrm{e}^{-d\cos\theta_{ijk}}\right]^{-1}
\end{aligned} \tag{9-20}
$$

这种势形式能够给出稳定的 Si 金刚石结构，并能广泛地应用于 Si 结构的计算。

2. Brenner 势

Brenner 势的两种形式也具有类似于 Tersoff 势的形式，它应用到 C 和碳化氢分子的研究。

（1）第一种 Brenner 势的形式

第一种 Brenner 势的形式如下

$$
\begin{aligned}
&U_{ijk} = f_c(r_{ij})\left[V_R(r_{ij}) + b_{ij}V_A(r_{ij})\right] \\
&V_R(r_{ij}) = \frac{D_e}{S-1}\exp\left[-\beta\sqrt{2S}\left(r_{ij}-r_c\right)\right] \\
&z_{ij} = \sum_{k\neq i,j} f_C(r_{ik})g(\theta_{ijk})\mathrm{e}^{m\left(r_{ij}-r_{ik}\right)} \\
&g(\theta) = \alpha\left\{1+\frac{c^2}{d^2}-\frac{c^2}{d^2+(h+\cos\theta)^2}\right\}
\end{aligned} \tag{9-21}
$$

它可以应用于计算金刚石和石墨的不同特性，也可应用于小的碳团簇的计算。

（2）第二种 Brenner 势的形式

第二种 Brenner 势的形式如下

$$
\begin{aligned}
&U_{ijk} = V_R(r_{ij}) - \bar{B}_{ij}V_A(r_{ij}) \\
&V_R(r_{ij}) = f_{ij}(r_{ij})\frac{D_{ij}}{S_{ij}-1}\exp\left[-\beta_{ij}\sqrt{2S_{ij}}\left(r_{ij}-r_e\right)\right] \\
&V_A(r_{ij}) = f_{ij}(r_{ij})\frac{S_{ij}D_{ij}}{S_{ij}-1}\exp\left[-\beta_{ij}\sqrt{2S_{ij}}\left(r_{ij}-r_e\right)\right] \\
&f_{ij}(r) = \begin{cases} 1, & r < r_e^{(1)} \\ \dfrac{1}{2}\left[1+\cos\left(\pi\left(r-r_e^{(1)}\right)\Big/\left(r_e^{(2)}-r_e^{(1)}\right)\right)\right], & r_e^{(1)} < r < r_e^{(2)} \\ 0, & r > r_e^{(2)} \end{cases} \\
&\bar{B}_{ij} = \frac{1}{2}\left(B_{ij}+B_{ji}\right) + F_{ij}(N_i^{(+)}, N_j^{(+)}, N_{ii}^{conj})
\end{aligned}
$$

$$B_{ij}=\left[1+G_{ij}+H_{ij}(N_i^{(H)},N_i^{(C)})\right]$$
$$G_{ij}=\sum_{k\neq i,j} f_{ik}(r_{ik})G_i(\theta_{ijk})\exp\left\{\alpha_{ijk}\left[\left(r_{ij}-r_e\right)-\left(r_{ik}-r_e\right)\right]\right\}$$
$$N_i^{(H)}=\sum_{j(=hydrogen)} f_{ik}(r_{ik})$$
$$N_i^{(C)}=\sum_{j(=carbon)} f_{ik}(r_{ik})$$
$$N_i^{(l)}=N_i^{(H)}+N_i^{(C)}$$
$$N_{ij}=1+\sum_{k\neq i,j} f_{ik}(x_{ik})+\sum_{l\neq i,j} f_{jl}(x_{jl})$$
$$f(x_{ik})=\begin{cases}1, & x_{ik}<2\\ \dfrac{1}{2}\left\{1+\cos\left[\pi\left(x_{ik}-2\right)\right]\right\}, & 2<x_{ik}<3\\ 0, & x_{ik}\geq 3\end{cases} \quad (9\text{-}22)$$
$$x_{ik}=N_k^{tot}-f_{ik}(r_{ik})$$
$$G_c(\theta)=a_0\left\{1+\frac{c_0^2}{d_0^2}-\frac{c_0^2}{d_0^2+(1+\cos\theta)^2}\right\}$$

式中，下标 e 均表示平衡状态时的值。

通过拟合不同碳氢分子的特性来确定上面的势参数。上面的势形式应用于计算许多碳氢分子原子逃逸能、不同系统在金刚石和石墨表面的吸附能等。

Tersoff 和 Brenner 势被广泛地应用于材料特性的计算，例如，微团簇的结构、晶格动力学、体积点缺陷、液态和无定形态、表面重构、表面融化以及 Si−C、Si−Ge 系统等。

9.3.2 Stillinger-Weber 势

Stilling−Weber 势是针对硅的性质而提出的一种包括两体和三体相互作用的经验势，被广泛应用于硅的体材料和表面的特性研究。

1. Stillinger-Weber 势原型

Stillinger−Weber 势是对势和三体势的组合，结合能 E 由下式给出

$$E=\frac{1}{2}\sum_{i<j=1} V_2(r_{ij})+\sum_{i<j<k=1} V_3(r_{ij},r_{ik},r_{jk})$$
$$V_2(r_{ij})=f_C(r)\left[A_1\phi_1(r)-A_2\phi_2(r)\right]$$
$$f_C(r)=\begin{cases}\exp\left[\dfrac{\mu}{r-R_e}\right], & if\quad r<R_e\\ 0, & otherwise\end{cases} \quad (9\text{-}23)$$
$$V_3(r_{ij},r_{ik},r_{jk})=Z\psi(r_{ij})\psi(r_{ik})g(\theta_i)$$
$$\psi(r)=\left[f_C(r)\right]^{\alpha}$$
$$g(\theta_i)=\left(\cos\theta_i-\cos\theta_0\right)^2$$

Stilinger−Weber 通过拟合金刚石结构的晶格常数、结合能，参照它的熔点得到势的参数，对液态 Si 的结构也能够较好地描述。Stillinger−Weber 势被广泛地应用于 Si 物理性质的研究。如研究团簇、晶格动力学、点缺陷、液态和无定形态、表面重构、界面等等。

2. Stillinger−Weber 势的修正形式

在计算 Si 的各种物理特性的过程中，出现了 Stilinger−Weber 势的很多修正形式，如 Feuston 和 Garofalini 在处理 SiO_2 系统时，对势部分采用了修正的 Born−Mayer−Huggins 离子势

$$V_2(r_i,r_j)=V_2(r_{ij})=A_{ij}\exp\left(\frac{r_{ij}}{\rho}\right)+\left(\frac{Z_iZ_je^2}{r_{ij}}\right)\mathrm{erfc}\left(\frac{r_{ij}}{\beta_{ij}}\right) \tag{9-24}$$

式中，Z_i 是离子电荷，ρ 和 β_{ij} 是可调参数。A_{ij} 由下式给出

$$A_{ij}=\left(1+\frac{Z_i}{n_i}+\frac{Z_j}{n_j}\right)\bullet b\exp\left[\frac{\left(\sigma_i+\sigma_j\right)}{\rho}\right] \tag{9-25}$$

其中，n_i 是价电子数，σ_i 是离子半径，b 是常数。

三体势部采用了修正的 SW 形式，其具体形式如下

$$V_3(r_i,r_j,r_k)=h(r_{ij},r_{ik},\theta_{jik})+h(r_{jk},r_{ji},\theta_{kji})+h(r_{ki},r_{kj},\theta_{ikj})$$

$$h(r_{ij},r_{ik},\theta_{jik})=\begin{cases}\lambda_i\exp\left[\dfrac{\gamma_i}{\left(r_{ij}-r_i^c\right)}+\dfrac{\gamma_i}{\left(r_{ik}-r_i^c\right)}\right]\left(\cos\theta_{jik}-\cos\theta_{jik}^c\right)^2, & \begin{matrix}r_{ij}<r_i^e\\ r_{ik}<r_i^e\end{matrix}\\ 0, & otherwise\end{cases} \tag{9-26}$$

式中，λ_i、γ_i、r_i^e 和 θ_{jik}^e 是常数。它们用这种势形式计算了 SiO_2 玻璃和（$Na_2O{\cdot}3SiO_2$）的结构，Si 的表面键缺陷结构的分子动力学模拟，Si 溶胶聚合过程的分子动力学模拟，惰性气体（N_2、Ar、Ne）在玻璃表面吸附的分子动力学模拟，Si 玻璃的表面驰豫等。

9.3.3 Slater−Kirkwood 类型的三体势

Bruch 和 McGee 在计算 4He 的三体聚合物的基态能量时，使用了 Slater−Kirkwood 模型表述附加的三体势，角度部分的函数形式采用了 Axilord−Teller 的势形式

$$V_3=\left\{-A\exp\left[-\alpha(r+s+t)\right]+C(rst)^{-3}\right\}\left(1+3\cos\theta_1\cos\theta_2\cos\theta_3\right) \tag{9-27}$$

式中，r、s 和 t 是三个原子组成的三角形的边长，θ_1、θ_2 和 θ_3 是三角形的三个内角。

这种三体势的形式广泛地应用于描述晶体中原子间的三体相互作用，并用来

计算材料的结构特性和动力学特性。

类似的势函数形式有 Bauer−Maysenholder−Seeger 势

$$U_{ijk} = Z\frac{\left(1+3\cos\theta_i\cos\theta_j\cos\theta_k\right)}{\left(r_{ij}r_{jk}r_{ki}\right)^m} \tag{9-28}$$

它用来描述 FCC 金属的点缺陷特性，它的参数通过拟合晶格常数、结合能和弹性常数来确定。

另一种类似的势函数形式 Pearson−Takai−Halicioglu−Tiler 势，被用来研究金刚石结构 Si 的相变、表面重构、表面点缺陷的形成和扩散等，它还应用于研究 GaAs、Si−GaAs、Au−GaAs 等二元和三元系统。

9.3.4 Chelikowsky−Philips−Kamal−Strauss 势

Chelikowsky−Philips−Kamal−Strauss 势被用来描写 Si 和 C

$$\begin{aligned}
&U_{ijk} = -\frac{a}{r_{ij}^2}\exp\left[-\beta_1 r_{ij}^2\right] - \frac{g_{ij}}{r_{ij}^2}\exp\left[-\beta_2 r_{ij}^2\right]\\
&g_{ij} = g_0 + g_1 S_{ij} S_{ij}\\
&S_{ij} = 1 + \left\langle\cos\left(3\theta_{ijk}\right)\right\rangle\\
&\left\langle f\left(\theta_{ijk}\right)\right\rangle = \left[f\left(\theta_{ijk}\right)\right]\big/[\mathrm{I}]\\
&\left[f\left(\theta_{ijk}\right)\right] = \sum_{k\neq i,j} f\left(\theta_{ijk}\right)\exp\left[-\left(\lambda_1\theta_{ijk}^4 + \lambda_2 r_{ijk}^4\right)\right]\\
&r_{ijk} = \frac{1}{2}\left(r_{ij} + r_{ik}\right)
\end{aligned} \tag{9-29}$$

这种形式的势通过拟合共价晶体到金属的转变获得势参数。利用它可以给出 Si 团簇的能量和结构，它也可以用来描述 C 和 C_{60}。

9.3.5 Khor−Das Sarma 势

Khor−Das Sarma 势的形式如下

$$\begin{aligned}
&U_{ijk} = Af_{ij}\left[e^{-\theta r_{ij}} - g_{ij}\left(1+h_{ij}\right)\right]\\
&f_{ij} = \mathrm{e}^{-\beta(r_{ij}-R_i)^\gamma}\\
&g_{ij} = \frac{B_0}{Z_i^\alpha}\mathrm{e}^{-\lambda r_{ij}}\\
&Z_i = \sum_j \exp\left[-\beta\left(r_{ij}-R_i\right)^\gamma\right]\\
&h_{ij} = \sum_{k\neq i,j}\left[\cos\eta(\theta_{ijk}-\theta_i) - 1\right]
\end{aligned} \tag{9-30}$$

其中，模型参数可以通过拟合金刚石结构的体模量和键的弯曲力常数来确定。该模型可以给出不同结构的结合能和平衡原子间距，它的改进形式应用到了研究 Si（111）表面的原子吸附和重构方面。

9.3.6　Biswas–Hamann 势

Biswas–Hamann 势的函数形式为

$$
\begin{aligned}
&V = V_2 + V_3 = \sum_{i<j} U_{ij} + \sum_{i<j<k} W_{ijk} \\
&U_{ij} = A_1 \mathrm{e}^{-\lambda_1 r_{ij}} + A_2 \mathrm{e}^{-\lambda_2 r_{ij}} \\
&W_{ijk} = \sum_{i=0}^{6} C_i \phi_i(r_{ij}) \phi_i(r_{ik}) P_i\left(\cos\theta_i\right) \\
&\phi_i(r) = \mathrm{e}^{-\alpha_i r}
\end{aligned}
\tag{9-31}
$$

这种势参数通过拟合体、表面、层以及裂纹结构的能量获得。它应用到 Si 的团簇和体的结构特性的研究。

9.3.7　Murrell–Mottram 势

Murrell–Mottram 势的多项式与 Biswas–Hamann 势的函数形式相同，只是 U_{ij} 和 W_{ijk} 的取法不同

$$
\begin{aligned}
&U_{ij} = D\left[1 + b\left(r_{ij} - r_e\right)\right] \mathrm{e}^{-b\left(r_{ij} - r_e\right)} \\
&W_{ijk} = P\left(Q_1, Q_2, Q_3\right) \mathrm{e}^{-\alpha Q_1}
\end{aligned}
\tag{9-32}
$$

式中，

$$
\begin{aligned}
P\left(Q_1, Q_2, Q_3\right) = {} & C_0 + C_1 Q_1 + C_2 Q_1^2 + C_3\left(Q_2^2 + Q_3^2\right) + C_4 Q_1^3 \\
& + C_5 Q_1\left(Q_2^2 + Q_3^2\right) + C_6\left(Q_3^3 - 3 Q_3 Q_2^2\right)
\end{aligned}
\tag{9-33}
$$

$$
\begin{bmatrix} Q_1 \\ Q_2 \\ Q_3 \end{bmatrix} = \begin{bmatrix} \dfrac{1}{\sqrt{3}} & \dfrac{1}{\sqrt{3}} & \dfrac{1}{\sqrt{3}} \\ 0 & \dfrac{1}{\sqrt{2}} & \dfrac{1}{\sqrt{2}} \\ \dfrac{\sqrt{2}}{\sqrt{3}} & -\dfrac{1}{\sqrt{6}} & -\dfrac{1}{\sqrt{6}} \end{bmatrix} \begin{bmatrix} r_{ij} - r_e \\ r_{ik} - r_e \\ r_{jk} - r_e \end{bmatrix}
\tag{9-34}
$$

对于 Si，这种势的参数通过拟合金刚石结构 Si 的结合能、晶格常数、弹性常数和它的 Raman 谱来确定。对于金刚石结构的 C 和 Si，这种势的形式在后续的研究中得到了发展和修正。

9.3.8 Erkoc 势

Erkoc 势采用 Biswas−Hamann 势的函数形式，其中 U_{ij} 和 W_{ijk} 的形式如下

$$U_{ij}=A\left\{\left(\frac{r_0}{r_{ij}}\right)^{2n}\exp\left[-2\alpha\left(\frac{r_{ij}}{r_0}\right)^2\right]-\left(\frac{r_0}{r_{ij}}\right)^{n}\exp\left[-2\alpha\left(\frac{r_{ij}}{r_0}\right)^2\right]\right\}$$

$$W_{ijk}=B\left(U_{ij}f_{ijk}+U_{ik}f_{ikj}+U_{jk}f_{jki}\right) \tag{9-35}$$

式中，$A=-4\varepsilon_0$，$\alpha=\ln(2)$，$n=\sqrt{\dfrac{r_0^2k}{2|\varepsilon_0|}-2\alpha}$，$f_{ijk}=e^{-\frac{r_{ij}^2+r_{ik}^2}{r_0^2}}$

这种势的形式被应用到研究微团簇的结构和能量。Erkoc 势的改进形式被应用到研究 FCC、B CC、HCP 等各种各样的金属、半导体等材料。

本节我们对一些典型的多体势函数进行了介绍。除此之外，还有很多的势函数模型应用到材料的各种特性研究。如：研究 Si 的小团簇、晶相、点缺陷和表面重构的 Bolding−Andersen 势；研究 Si 的声子谱、结合能、表面重构的 Brenner−Garrison 势；研究 Si 团簇和表面的 Niistriotis−Flytzanis−Farantos 势；研究硅酸盐（Mg_2SiO_4）的 Price−Parker−Leslie 势等。

9.4 嵌入（埋入）原子势（Embedded−atom method，EAM）

1983 年，Daw 和 Baskes 用半经验的量子力学的方法计算氢脆问题时，提出了计算原子间相互作用的埋入原子法，是在电子的密度泛函理论和有效介质理论的基础上发展起来的一种构造原子间相互作用势的方法。其基本思想是金属中的离子间相互吸引可视为将离子核嵌入到自由电子中引起电子气体能的降低所对应的结合能。Foiles 等人对金属 EAM 势的构造做了重要改进，并给出了 Au、Ag、Cu、Pt、Pd、Ni 等金属的 EAM 势，使之不仅适用于纯金属材料的研究，而且可以对合金材料给出比较理想的描述。人们根据 EAM 理论，先后构造出适用于不同材料的多种形式的 EAM 势。EAM 方法在解决固体声子谱、液态金属、缺陷、合金、杂质、断裂、表面结构、表面吸附、表面迁移、表面有序−无序相变、表面有序合金、表面声子、团簇等诸多领域均取得了巨大成功。

9.4.1 原子对势理论与准服子理论和有效介质理论

由 Daw 和 Baskes 所提出的嵌入原子方法是基于准原子理论和有效介质理论建立起来的。因此在介绍和叙述 EAM 理论之前，简要地叙述一下原子对势和准原子理论。

1. 原子对势理论

对于纯晶体的总能量，在未经形变时，按照原子对势理论应为 $E_b = 1/2\sum V(r)$，即两体势。但两体势有固有的缺陷。

2. 准原子理论和有效介质理论

1980 年，Stott 和 Zarerriba 在密度函数理论的基础上，提出了准原子理论。首先，他们把杂质离子与其电子屏蔽云看做一个整体，称为一个准原子。对于系统的总能，按照准原子理论应为 $E = E_x\left(\rho_n(r)\right) + E_{latt}$。其中，第一项为准原子的能量，第 2 项为晶格驰豫能。

与此同时，由 Norskov 和 Lang 提出了类似的理论：有效介质理论，其基本想法是将能量近似地用嵌入一个原子于均匀电子气中，该嵌入原子所在处的基体电子密度的函数表示。在此，明确地提出了嵌入能量的观念，但是这仍然不能回避定义基体体积的问题，因此就无法解决断裂和氢脆这些问题。

9.4.2　原型 EAM 理论

针对原子对势理论及准原子理论存在的缺陷，20 世纪 80 年代中期，Daw 和 Baskes 首先以解决氢脆和断裂为目的，提出了嵌入原子方法（EAM）理论。嵌入原子方法模型的发展源于密度泛函理论，即认为原子系统的能量可由其电子密度的函数精确给出，这个函数的准确形式也许无法知晓，但可通过研究分子或晶体的电子结构而近似得到。鉴于此就可认为在原子系统中嵌入一个原子所引起的能量变化由嵌入前系统电子密度的函数给出，如果能近似得到这个嵌入函数，则可对系统中每个原子进行嵌入处理而近似计算系统能量。零级近似下，嵌入能等于将原子嵌入到均匀电子气的能量，加上经典的原子相互作用就能得到系统能量。均匀电子气的嵌入能可由第一原理计算得到，这就为 EAM 模型的发展提供了理论基础。Norskov 用这一方法计算了氢和氮杂质在各种金属中的能量，与实验结果定性相符。

Daw 和 Baskes 进一步提出了两点改进：

1）在能量表达式中不用第一原理计算，而采用易于参数化的原子相互作用势 $V(r)$和嵌入能 $F(\rho_i)$计算，$F(\rho)$是原子 i 处的总电子密度 ρ_i 的函数。

2）对固体中的每一个原子分别计算这两项能量，对所有原子求和获得系统总能量 E_{tot}，即

$$
\begin{aligned}
E_{tot} &= \sum_{i=1}^{n} E_i \\
E_i &= \frac{1}{2}\sum_{j} V(r) + F(\rho_i) \\
\rho_i &= \sum_{j\neq i} \phi\left(r_{ij}\right)
\end{aligned}
\tag{9-36}
$$

其中，原子的电子密度$\phi(r)$可由 Hartree−Fock 近似计算，$F(\rho)$则通过拟合实验数据优化得到，从而成功获得了用途广泛的半经验近似方法。随后，Manninen、Jacobsen 等进一步从理论上讨论了 EAM 与密度泛函方法的联系，使这一半经验模型建立在坚实的理论基础上。

EAM 理论克服了对势理论以及准原子理论的固有缺陷。因为若是只考虑对势，必然导致弹性常数 $C_{12} = C_{44}$ 的错误结论；若是像准原子理论那样，只考虑嵌入能，则得到 $C_{44} = 0$ 的结论，这与立方晶体的实际情况相违背。EAM 理论不仅克服了对势理论人为加上体积相关能量项，因而无法准确定义体积的困难，也避免了准原子理论的问题。建立起了描述固体的原子层次理论。它已经成为计算材料科学原子层次理论的代表性理论。

EAM 模型的基本思想是：把系统中的每一个原子都看成是嵌入在由其他原子组成的基体中的杂质，将系统的能量表示为嵌入能和相互作用势能之和，从而将多原子相互作用归结于嵌入能，对嵌入能的计算作了两个假设，其一是假设嵌入能是局域电子密度及其高阶导数的函数；其二是假设固体的电子密度可表示为原子电子密度的线性叠加，并假设原子的电子密度呈球对称分布。这就导致了与实际情况的偏差，特别是对有角度分布的定向成键系统产生较大误差。同时金属键的特性要求嵌入函数的曲率为正，这就无法描述负 Cauchy 压的金属元素及其合金。至于如何确定嵌入函数和势函数，在 Daw 和 Baskes 最初的模式中，所有的计算都采用数值拟合方法，未给出具体函数，经 Foiles 等人发展后，选定了两种势的函数形式，通过拟合 Rose 方程而得到嵌入函数，其参数的确定还是通过数值计算逐步逼近得到：为了解决负 Cauchy 压元素的问题，Baskes 考虑了定向键效应，并引入有角度分布因子的修正模型，在此模型中，是先设定嵌入函数，再通过拟合 Rose 方程确定两体势，模型参数同样由数值拟合得到。由于模型参数的确定不是分析形式的，不能直接和具体的物理参数相联系，只适应所处理的特定系统，如对元素计算的参数不能推广到其合金系统，对合金系的计算则需用合金特性参数进行似合，这就使 EAM 理论的应用受到了很大限制，特别不适合用来系统地处理有关合金方面的问题。

在 EAM 模型原框架中，电子密度球对称分布假设在一些情形下已严重偏离实际情况，如 d 电子轨道未满的过渡族元素，金刚石结构的半导体元素及轨道杂化的合金系统等，对总能量的计算有很大影响。同时嵌入函数的正曲率要求也无法处理像 Cr、Cs 等具有 Cauchy 负压的元素及合金，因此，适当的修正很有必要。

Baskes 等人对原有的 EAM 理论的修正是保持原理论框架不变，针对实际的原子电子密度偏离球对称分布的情况，在基体电子密度求和中引入原子电子密度分布的角度依赖因素，没有考虑修正基体电子密度是原子电子密度的线性叠加这一假设。经过修正后，Baskes 理论已能描述负 Cauchy 压的金属元素，但确定参数

的过程同样采用数值拟合，不是分析型的，相当繁杂，应用上仍受很大的限制，尤其是应用到合金系统上更为复杂，因而具体应用很鲜见。1994 年 Baskes 等在谈到他们修正过的 EAM 理论应用至合金形成的计算时曾指出，这方面需要进行很多的工作，然而，至今仍未见他们自己在这方面有结果报告，足见用该 EAM 理论描述合金之困难。

Pasianot 等试图在总能量中加入一修正项。以修正原 EAM 模型中两条基本假设所引起的能量差。Pasianot 等的“嵌入缺陷”模型是在总能量表达式中引入一与键角有关的多体剪切修正项，具体算法类似于 Kanzaki 偶极张量的计算，但未能给出一个明确的修正项表达式，这一模型可成功地描述 Nb、Fe、Cr 等元素，遗憾的是未能进一步发展和应用，主要原因在于模型不是分析型的，应用起来很繁杂。

9.4.3　分析型嵌入原子（AEAM）理论

为了完全确定对势函数、电子密度以及嵌入函数，有三个模型参数必须确定。由于 EAM 理论中没有得出这些模型参数与元素体性能之间的分析表达式，因而使两体势和嵌入函数没有分析表达式，未能直接和具体的物理参数相联系，只能用数字解法，因而无法使原型 EAM 理论系统化地推广和应用，于是产生了分析型嵌入原子（AEAM）理论。

Johnson 在 Foiles 和 Baskes 等人的基础上，经验地给出了电子密度、两体势和嵌入能的分析函数形式，基于弹性常数与能量的关系，体系平衡条件和形成空位的物理过程，建立起模型参数和物理参数对应关系的解析表达式，通过拟合元素的物理性能参数确定各分析函数的模型参数，得到特定结构金属元素的分析型 EAM 模型，通过所构建的合金势计算合金的物理性能。使特定结构的金属及合金系统的 EAM 模型初步普适化。

对于嵌入函数 $F(\rho)$、两体势 $V(r)$和原子的电子密度分布函数$\phi(r)$的确定，Johnson 对特定结构金属元素先设定这些函数的具体形式，通过拟合金属的结合能、弹性常数、单空位形成能来确定模型参数，从而建立起一系列的模型参数与物理参数相联系的分析表达式。

Johnson 推导出了一个普适的嵌入函数形式

$$F(\rho) = -F_0\left[1 - n\ln\left(\frac{\rho}{\rho_e}\right)\right]\left(\frac{\rho}{\rho_e}\right)^n \tag{9-37}$$

并且，针对不同结构的金属势函数和电子密度函数采用不同的函数形式。式（9-37）中下标 e 表示平衡态，下同。

面心立方（fcc）和密排六方（hcp）金属，其势函数和电子密度函数为

$$V(r)=V_e\exp\left[-\gamma\left(\frac{r}{r_e}-1\right)\right]$$
$$\phi(r)=\phi_e\exp\left[-\beta\left(\frac{r}{r_e}-1\right)\right] \tag{9-38}$$

体心立方（bcc）金属，其势函数和电子密度函数为

$$V(r)=k_0+k_1\left(\frac{r}{r_e}-1\right)+k_2\left(\frac{r}{r_e}-1\right)^2+k_3\left(\frac{r}{r_e}-1\right)^3$$
$$\phi(r)=\phi_e\left(\frac{r_e}{r}\right)^\beta \tag{9-39}$$

所有模型参数可由其与物理性能参数对应关系的解析表达式计算出。尤为重要的是，由于构建了分析型的合金势，对纯金属计算的模型参数同样适应于合金系统的计算，其合金势函数的计算为（对于 A−B 二元合金）

$$V^{ab}(r)=\frac{1}{2}\left[\frac{\phi^b(r)}{\phi^a(r)}V^{aa}(r)+\frac{\phi^a(r)}{\phi^b(r)}V^{bb}(r)\right] \tag{9-40}$$

式中，$V^{aa}(r)$和 $V^{bb}(r)$分别表示 A 原子核 B 原子的两体势；$\phi^a(r)$和$\phi^b(r)$分别表示 A 原子和 B 原子的电子密度分布函数。

由此可讨论合金相的稳定性，并计算了 fcc 二元合金的稀溶解热及形成焓。

9.4.4 改进分析型 EAM 模型（MAEAM）

张邦维等人综合考虑以上各种 EAM 模型的优缺点后，一方面按照 Pasianot 的作法，在原 EAM 理论的基本表达式，即系统总能量表达式中经验地加入一修正项，用以描述两条基本假设所引起的能量偏离，并且具体给出了该修正项的分析表达式。另一方面，他们发展了一个能较好描述各种结构金属的原子间势函数的分析表达式，对于嵌入函数、原子电子密度和合金势，则仍然沿用 Johnson 的原有分析表达式。这样，他们所提出的分析型 EAM 理论对 fcc，bcc 和 hcp 各种结构的金属皆能进行描述，且能处理负 Cauchy 压的元素，成为一个普适分析型 EAM 模型。

在张邦维等人的普适分析型 EAM 理论中，对式（9-36）进行了修正，具体形式为

$$E_i=\frac{1}{2}\sum V(r_{ij})+F(\rho_i)+M(P_i)$$
$$P_i=\sum_{j\neq i}\phi^2\left(r_{ij}\right) \tag{9-41}$$

修正项 $M(P)$的分析表达式为

$$M(P)=\alpha\left(\frac{P}{P_e}-1\right)^2\exp\left[-\left(\frac{P}{P_e}-1\right)^2\right] \tag{9-42}$$

两体势的分析表达式为

$$V(r)=k_0+k_1\left(\frac{r}{r_e}\right)^2+k_2\left(\frac{r}{r_e}\right)^4+k_3\left(\frac{r_e}{r}\right)^{12} \tag{9-43}$$

其中，下标 e 均表示平衡状态时的值。至于嵌入函数 $F(\rho)$、原子电子密度$\phi(r)$和合金势 $V^{ab}(r)$则仍沿用 Johnson 所提出的分析形式表达式（9-37）～（9-40）。所有模型参数同样通过拟合纯金属的结合能、原子体积、体积模型、剪切模量和单空位形成能来确定。

此模型可很好地描述包括 Cr 在内的 bcc 金属，包括 Cs 在内的碱金属较好地描述 hcp 金属和 fcc 金属，由所确定的模型参数计算的弹性常数和空位形成能与实验测定结果符合较好。对 bcc 金属计算的单空位、双空位的形成能和迁移能与实验结果也很好相符。计算的各类金属间的二元合金稀溶解热和全成分范围的形成热也和实验结果基本相符。

胡望宇等在张邦维分析型 EAM 模型的基础上作了如下两点改进。

1）重新构造了金属的对势函数，并修改了原模型中的修正项的表达式，使该函数的函数值、一阶导数值和二阶导数值都比较小，以利于分子动力学的计算。

2）原模型中假定 hcp 金属均具有理想轴比，只用 3 个弹性常数拟合计算，对具体结构稳定性和空位移动能的计算与实验结果相差甚大。考虑到 hcp 金属的具体结构并运用全部 5 个弹性常数进行计算，重新构造了一个 hcp 分析型的 EAM 模型。

运用此模型计算了 hcp 金属的物性（包括声子谱、比热容、热膨胀等）、计算了晶体的点缺陷及扩散特性、层错和合金稀溶解热与形成热，计算结果与实验结果符合较好。

以上说明这种经验性的修正方法，尽管其物理意义有待进一步明确，具体函数形式有待进一步改善，却能直接明了地给出其分析模型，很有发展前途。

1. 金属元素的 MAEAM 模型

本书所采用的改进分析型 EAM 模型（MAEAM），嵌入一个原子引起的能量变化仍采用式（9-39）的形式。考虑到计算上的方便，以嵌入原子为坐标原点，将能量表达式简化为

$$E=F(\rho)+\frac{1}{2}\sum_m V(r_m)+M(P) \tag{9-44}$$

式中，$\rho=\sum_m \phi\left(r_m\right)$，$\phi(r_m)=\phi_e\left(r_1/r\right)^6$，$m$ 代表第 m 近邻原子。其中的嵌入函数、原子电子密度及电子分布函数仍沿用 Johnson 的形式。

F_0 的确定依据平衡条件下的能量关系得到

$$F_0=E_c-E_{1f} \tag{9-45}$$

修正项 $M(P)$的分析表达式为

$$M(P)=\alpha\left\{1-\exp\left[-\left(\ln\left|\frac{P}{P_e}\right|\right)^2\right]\right\} \tag{9-46}$$

对于 bcc，fcc 金属的参量 P 表示为

$$P=\sum_m \phi^2\left(r_m\right) \tag{9-47}$$

对于 hcp 金属的参量 P 表示为

$$P=\sum_m \phi^2\left(r_m\right)\frac{r_{mx}^2+r_{my}^2+r_{mz}^2}{r_m^2} \tag{9-48}$$

式中，r_{mx}、r_{my} 和 r_{mz} 分别表示与原点原子相互作用的原子处于 r_m 位置的各坐标分量。

对于势函数，在 Johnson 的模型中，是用一个三次多项式表示的，并在势函数两端分别进行挺坚和截尾处理，通过引入一个新的势函数多项式，可以不必进行挺坚处理。针对不同的晶体结构采用了不同的势函数形式。

（1）bcc 金属的对势函数

Johnson 和 Oh 最早提出 bcc 金属的分析型 EAM 模型，他们详细地阐述了构建分析型 EAM 模型的基本思路。张邦维等运用此模型对过渡族 bcc 金属的稀溶解热和形成热进行了计算，发现大多数合金系统的计算值与实验结果或 Miedema 的计算值是基本相符的，但 Ta−W、W−M 合金系的计算结果与实验结果相反，此模型还存在一个严重缺陷，就是不能处理负 Cauchy 压的元素与合金，为此张邦维对势函数进行了改进并加入一能量修正项，解决了此问题，并在后来的一篇论文中对这一改进分析型 EAM 模型（MAEAM）作了简单综述。

对于 bcc 金属，可采用如下对势函数式

$$V(r)=\sum_{j=-1}^{j=4}k_j\left(\frac{r}{r_1}\right)^j \tag{9-49}$$

在现在的模型中，原子间相互作用计算时只计及最近邻与次近邻的原子，$V(r)$函数在第二近邻与第三近邻之间截尾。截尾距离 $r_{ce}=r_2+k_{ce}\left(r_3-r_2\right)$，其中 k_{ce} 是一调节参数，通过拟合 bcc 金属的物理性质来确定。在该截尾处，势函数值和势函数的一阶导数值均为零，即：$V(r_{ce})=0,\ \ V'(r_{ce})=0$ 。

因此，可得到 bcc 金属的完整势函数为

$$V(r)=\begin{cases}\displaystyle\sum_{j=-1}^{j=4}k_j\left(\frac{r}{r_1}\right)^j & r\leqslant r_{ce}\\ 0 & r>r_{ce}\end{cases} \tag{9-50}$$

在计算结合能与原子体积关系时，为避免结合能不连续的问题，电子密度分

布函数与势函数分别采用了不同的截尾处理。电子密度分布函数计算到第三近邻的原子，在第三近邻与第四近邻的四分之三处截尾。截尾距离 $r_{cf}=r_3+k_{cf}\left(r_4-r_3\right)$，其中 k_{cf} 取为 0.75。在该截尾处，函数值和函数的一阶导数值均为零，即：$\phi(r_{cf})=0,\ \phi'(r_{cf})=0$。

势函数中的模型参数可以通过空位形成能、晶格平衡条件和弹性常数等物理性质来确定。

元素的未弛豫的空位形成能可近似用对势表示如下

$$E_{1f}=-\frac{1}{2}\sum_m V(r_m) \tag{9-51}$$

在晶格没有应力时，晶体的平衡条件可以表示如下

$$\frac{1}{2}\sum_m \frac{r_{mi}r_{mj}}{r_m}V'(r_m)=0 \tag{9-52}$$

bcc 金属元素的弹性常数 C_{ijkl}，可用下式计算

$$\begin{aligned}\Omega_0 C_{ijkl}&=F''(\rho)\sum_m \frac{r_{mi}r_{mj}}{r_m}\phi'(r_m)\sum_m \frac{r_{mk}r_{ml}}{r_m}\phi'(r_m)\\&+\frac{1}{2}\sum_m \frac{r_{mi}r_{mj}r_{mk}r_{ml}}{r_m}\left[V''(r_m)-\frac{V'(r_m)}{r_m}\right]\\&+4M''(P)\sum_m \frac{r_{mi}r_{mj}}{r_m}\phi(r_m)\phi'(r_m)\sum_m \frac{r_{mk}r_{ml}}{r_m}\phi(r_m)\phi'(r_m)\end{aligned} \tag{9-53}$$

有三个独立的弹性常数，也就有三个方程。联立三个弹性常数方程和一个平衡条件方程，两个截尾条件方程和一个空位形成能方程就可以确定 α 和 k_j（j=−1，0，1，2，3，4）7 个模型参数。

根据收集的有关资料，整理了过渡组 bcc 金属元素的输入参数（列于表 9-1 中），舒小林根据表 9-1 中的 bcc 金属元素输入参数，计算得到了该模型参数（见表 9-2）。

表 9-1　bcc 金属元素的输入参数

元素	晶格常数 a/nm	结合能 E_c/eV	空位形成能 E_{1f}/eV	C_{11}/GPa	C_{12}/GPa	C_{44}/GPa
Cr	0.28846	4.10	1.80	346	66.0	100
Fe	0.28664	4.28	1.79	230	135	117
Mo	0.31468	6.82	3.10	459	168	111
Nb	0.33007	7.57	2.75	245	132	28.4
Ta	0.33026	8.10	2.95	262	156	82.6
V	0.30282	5.31	2.10	230	120	43.2
W	0.31650	8.90	3.95	517	203	157

表 9-2 bcc 金属元素的模型参数

参数	Cr	Fe	Mo	Nb	Ta	V	W
α	−0.3557	0.0032	0.0475	0.2621	0.1157	0.1372	0.0621
n	0.98	0.35	0.52	0.41	0.47	0.42	0.36
k_c	0.350	0.510	0.394	0.600	0.550	0.600	0.410
F_0	2.30	2.49	3.72	4.82	5.15	3.21	4.95
k_1/eV	8547.89	1478.54	8859.23	1723.79	3598.77	673.039	9973.49
k_2/eV	−7769.6	−1287.3	−8045.3	−1564.0	−3149.4	−629.26	−8992.5
k_3/eV	3522.52	563.579	3644.56	705.897	1375.68	293.740	4047.12
k_4/eV	−636.70	−98.957	−658.07	−126.45	−239.53	−54.434	−726.38
k_{-1}/eV	1030.26	199.603	1071.46	208.294	469.637	78.5707	1226.04
k_0/eV	−4694.6	−855.74	−4872.4	−947.97	−2055.6	−361.96	−5528.4

（2）fcc 金属元素的对势函数

fcc 金属分析型 EAM 模型最初由 Johnson 在 1988 年提出，当时只考虑最近邻原子的相互作用，给出了建立分析型 EAM 模型的基本思路。随后，Johnson 建立了具体的计算方法并考虑了更远近邻原子的相互作用。为了体现模型的应用价值，他们用所建模型计算了几种 fcc 金属形成二元合金的形成热，计算结果和实验结果及第一原理的计算值大致相符。由于 Johnson 模型不能处理负 Cauchy 压元素，张邦维等对 Johnson 模型进行改进，试图发展一种普适分析型的 EAM 模型，并据此重新计算了 fcc 金属二元合金的形成热，计算时考虑了最近邻和次近邻原子的相互作用，采用和 bcc 金属完全相同的势函数，所得结果比 Johnson 的略有改善。运用此模型处理 fcc 金属的扩散问题时，发现计算的空位迁移能大大小于实验结果，个别元素甚至出现负的空位迁移能，说明模型在处理元素特性问题时就不能自洽。胡望宇等在其模型中对势函数和模型方程进行了调整，并考虑到第五近邻原子的相互作用，解决了元素特性自洽性问题。

fcc 金属采用了如下形式的势函数

$$V(r)=k_0+k_1\left(\frac{r}{r_{1e}}\right)+k_2\left(\frac{r}{r_{1e}}\right)^2+k_3\left(\frac{r}{r_{1e}}\right)^6+k_4\left(\frac{r}{r_{1e}}\right)^{-12}+k_5\left(\frac{r}{r_{1e}}\right)^{-1} \tag{9-54}$$

在通常使用的模型中，fcc 金属原子间相互作用计算时计算到第五近邻的原子，$V(r)$ 函数在第五近邻与第六近邻之间截尾。截尾距离 $r_{ce}=r_5+k_{ce}\left(r_6-r_5\right)$，其中 k_{ce} 是一调节参数，通过拟合 fcc 金属的物理性质来确定。在该截尾处，势函数值和势函数的一阶导数值均为零，即：$V(r_{ce})=0,\ V'(r_{ce})=0$。

在计算结合能与原子体积关系时，为避免结合能不连续的问题，fcc 金属的电子密度分布函数仍采用了不同于势函数的截尾处理方式。电子密度分布函数 $\phi(r)$ 计算到第六近邻的原子，在第六近邻与第七近邻的四分之三处截尾。截尾距离

$r_{cf}=r_6+k_{cf}\left(r_7-r_6\right)$，其中 k_{cf} 取为 0.75。在该截尾处，函数值和函数的一阶导数值均为零，即：$\phi(r_{cf})=0,\ \phi'(r_{cf})=0$。

fcc 金属的势函数中模型参数同样可以通过空位形成能、晶格平衡条件和弹性常数等物理性质来确定。

对于电子密度 $\phi(r)$ 及参数 F_0 的计算方法和 bcc 金属的完全相同，但 fcc 金属的原子体积 $\Omega_0=a^3/4$（a 为晶格常数）。

于是可得到 fcc 金属的完整势函数

$$V(r)=\begin{cases}k_0+k_1\left(\dfrac{r}{r_{1e}}\right)+k_2\left(\dfrac{r}{r_{1e}}\right)^2+k_3\left(\dfrac{r}{r_{1e}}\right)^6+k_4\left(\dfrac{r}{r_{1e}}\right)^{-12}+k_5\left(\dfrac{r}{r_{1e}}\right)^{-1}, & r\leqslant r_{ce}\\ 0, & r>r_{ce}\end{cases} \tag{9-55}$$

表 9-3 为根据收集的有关资料整理的过渡元素 Ag、Al、Au、Cu、Ir、Ni、Pb、Pt 和 Rh 9 个 fcc 金属的输入参数，表 9-4 为舒小林根据表 9-3 中的数据对 fcc 金属元素计算得到的模型参数。

表 9-3　fcc 金属的输入参数

元素	晶格常数 a/nm	结合能 E_c/eV	空位形成能 E_{1f}/eV	C_{11}/GPa	C_{12}/GPa	C_{44}/GPa
Ag	0.40857	2.95	1.10	123	92.0	45.2
Al	0.40496	3.39	0.64	116.3	64.8	30.9
Au	0.40788	3.81	0.90	190	161	42.3
Cu	0.36147	3.49	1.17	169	122	75.3
Ir	0.38389	6.94	3.50	600	260	270
Ni	0.35236	4.44	1.45	247	153	122
Pb	0.38907	3.89	1.30	224	173	71.6
Pt	0.39239	5.84	1.20	347	251	76.5
Rh	0.38401	5.75	2.90	412.6	193.5	184.1

表 9-4　fcc 金属的模型参数

参数	Ag	Al	Au	Cu	Ir	Ni	Pb	Pt	Rh
α	−0.00016	0.00018	0.0048	0.0032	−0.3641	0.0035	0.00093	0.0051	−0.1572
n	0.82	0.563	0.94	0.60	0.77	0.42	0.96	0.98	0.6
k_c	0.3	0.1	0.3	0.3	0.1	0.1	0.3	0.3	0.1
F_0	1.85	2.75	2.91	2.32	3.44	2.99	2.59	4.64	2.85
k_0/eV	0.2822	−1.1175	1.1566	1.4995	26.5881	3.3720	1.7549	−0.5866	−18.347
k_1/eV	0.1475	1.2149	−0.4510	−0.6502	−14.4873	−1.5687	−0.6746	1.4604	−10.132
k_2/eV	−0.0803	−0.3608	0.0541	0.0977	2.6168	0.2435	0.0767	−0.5196	1.8659
k_3/eV	0.0030	0.0012	0.00006	−0.00006	−0.0013	−0.00001	0.0001	0.0020	−0.0012
k_4/eV	0.0847	0.0502	0.0803	0.0988	0.4702	−0.1520	0.1185	0.1345	0.3209
k_5/eV	−0.5549	0.1116	−0.9523	−1.1783	−15.9145	−2.4149	−1.4438	−0.7491	−10.887

（3）hcp 金属的对势函数

对立方晶体结构的 bcc 和 fcc 元素，无论是 FS 还是 EAM 多体势都取得了较大进展，对这些金属组成的合金系统的原子模拟也相当成功，如对镍铝合金系的物理及力学性能的预测。相对而言，对 hcp 金属多体势的研究进展就甚为缓慢。正如 Bacon 教授对 hcp 金属计算机模拟的综述文章所言，hcp 金属晶体缺陷的模拟计算大多还是沿用两体势，且只能近似描述 c/a 比接近理想轴比的元素，尽管已开始发展 hcp 元素的多体势，但能满足具体模拟计算多项判据要求的多体势还远未构成，因此，构建 hcp 元素完备的多体势就成了发展原子尺度材料设计理论的当务之急。

20 世纪 80 年代末期，Johnson 首先构筑了 hcp 元素（Mg、Ti、Zr）的分析型 EAM 多体势，只采用一个平衡条件，通过拟合结合能、空位形成能、平均体模量和剪切模量来确定模型参数，由此模型计算的 5 个弹性常数与实验值符合较好，但 c/a 比却接近理想轴比，因此偏离理想轴比较大的元素就不能采用此模型计算。为了发展更普遍适应的 hcp 元素的多体势，Igarashi 和 Vitek 等人构筑了 Be、Hf、Ti、Ru、Zr、Co、Mg、Zn 8 个 hcp 元素的 FS 多体势，这是截至目前为止处理 hcp 元素最成功的多体势，计算了层错能、声子谱和自间隙原子特性，但由于没有拟合原子平衡距离之间的任何信息，使得所构多体势非常硬，计算的自间隙原子形成能特别大，因此在原子模拟计算中没有被采用，而且不能应用于合金系的计算。Pasianot 认为这些多体势的问题在于没有考虑晶格内禀自由度对弹性常数的贡献，因而就构筑了 Hf、Ti、Mg、Co 的 EAM 多体势，但他仅计算了这些元素的层错能，对空位、间隙原子等未作计算，模型的合理性不得而知，也就没引起人们的注意。关于 hcp 金属 EAM 多体势处理元素最多的是 Johnson 和 Baskes，在 1994 年发表的论文构筑了 18 个 hcp 元素的多体势，由于采用 Baskes 定向键修正思想，使得计算过程非常复杂。他们仅计算了空位、层错和表面特性，计算的双空位结合能与实验结果完全相反，且未对自间隙原子特性作任何报道，他们还允诺发表应用于 hcp 金属合金系统上的结果，但至今未见发表出来，说明模型存在较大问题。为了得到实用的 hcp 金属多体势，Ackland 进行了简化处理，在拟合三个弹性常数和一个晶格常数基础上构筑了关于 Ti 和 Zr 的 FS 多体势，并应用于核辐照联级碰撞中晶体缺陷形成过程的模拟计算，较好地解释了相关实验现象，但模型过于简化，只能处理接近理想 c/a 比的元素，不能推广到一般的 hcp 元素。

从以上事实可以看到，关于 hcp 元素多体势的研究引起了国际学术界的普遍关注，是材料设计领域的一个国际难题，同时 hcp 金属元素中包括了许多极有应用价值的元素，如 Ti、Mg、Hf、Co、Be、Zn 是重要的有色金属，Zr、U 是核材料中不可缺少的元素，Nd、Pr、Tb、Y、Dy、Er、Ho、Gd 等稀土元素在稀土合金中应用广泛，Os、Re 是高密度合金的重要组元等。发展 hcp 元素的多体势并开

展有关性能或过程的模拟计算，将会极大地促进有色金属合金和稀土合金的开发和应用。因此，发展系统而完备的 hcp 元素的多体势并进行原子模拟和性能预测不仅具有重要的科学意义，也具有难以估量的潜在应用价值。

hcp 金属的结构与立方晶体相比有显著差异，对于立方晶体只有 7 个模型方程，即 3 个弹性常数方程、1 个平衡方程、1 个空位形成能方程和两个势函数截尾条件方程，可确定 7 个模型参数，其中势函数的 6 个参数和修正函数中的一个参数。在 hcp 金属中，有 5 个独立的弹性常数，即 5 个弹性常数方程，由于 a 轴与 c 轴的各向异性存在 2 个平衡方程，1 个空位形成能方程，再加上 2 个势函数截尾条件方程，共有 10 个模型方程。为了描述电子密度高次项的各向异性效应，在修正函数中须引入 1 个各向异性修正因子，因而修正函数中包含 2 个模型参数，这与 hcp 金属存在 2 个 Cauchy 关系相对应，其余 8 个模型参数放在势函数中。对于 hcp 金属所考虑的近邻数也有较大差异，通过对比计算发现计算到原子第七近邻相互作用结果较为理想。

势函数的形式为

$$V(r)=\sum_{j=-1}^{j=6}k_j\left(\frac{r}{r_1}\right)^j \tag{9-56}$$

对电子密度高次项贡献因子 P 的计算与立方晶体不同，对于立方晶体，只有一个 Cauchy 关系，原子的排列呈现立方对称，因而修正函数只包含一个模型参数。对于 hcp 金属，存在两个 Cauchy 关系，且原子排列呈现各向异性，为了计入这种各向异性，在计算 P 时就引入一个各向异性因子，对于参考原子位于坐标原点（0，0，0）时的形式为

$$P=\sum_m \phi^2\left(r_m\right)\frac{r_{mx}^2+r_{my}^2+\beta r_{mz}^2}{r_m^2} \tag{9-57}$$

式中，r_{mx}、r_{my} 和 r_{mz} 分别表示与原点原子相互作用的原子处于 r_m 位置的各坐标分量。

在本书使用的模型中，hcp 金属原子间相互作用计算时计算到第七近邻的原子，$V(r)$函数在第七近邻与第八近邻之间截尾。截尾距离 $r_{ce}=r_7+k_{ce}\left(r_8-r_7\right)$，其中 k_{ce} 是一调节参数，通过拟合 hcp 金属的物理性质来确定。在该截尾处，势函数值和势函数的一阶导数值均为零，即：$V(r_{ce})=0,\ V'(r_{ce})=0$ 。

在计算结合能与原子体积关系时，为避免结合能不连续的问题，hcp 金属的电子密度分布函数仍采用了不同于势函数的截尾处理方式。电子密度分布函数$\phi(r)$计算到第八近邻的原子，在第八近邻与第九近邻的四分之三处截尾。截尾距离 $r_{cf}=r_8+k_{cf}\left(r_9-r_8\right)$，其中 k_{cf} 取为 0.75。在该截尾处，函数值和函数的一阶导数值均为零，即：$\phi(r_{cf})=0,\ \phi'(r_{cf})=0$ 。

hcp 金属的势函数中模型参数同样可以通过空位形成能、晶格平衡条件和弹性常数等物理性质来确定。

于是可得到 hcp 金属的完整势函数

$$V(r)=\begin{cases}\sum_{j=-1}^{j=6}k_j\left(\frac{r}{r_1}\right)^j & r\leqslant r_{ce}\\ 0 & r>r_{ce}\end{cases} \tag{9-58}$$

hcp 元素的未弛豫的空位形成能的对势形式与 bcc 元素相同。

由于 P 的变化将会引起弹性常数 C_{ijkl} 的变化，因此，hcp 元素 C_{ijkl} 与 bcc 元素的有所不同，其计算公式如下

$$\begin{aligned}\Omega_0 C_{ijkl} &= F''(\rho_e)\sum_m \frac{r_{mei}r_{mej}}{r_{me}}\phi'(r_{me})\sum_m \frac{r_{mek}r_{mel}}{r_{me}}\phi'(r_{me})\\ &+\frac{1}{2}\sum_m \frac{r_{mei}r_{mej}r_{mek}r_{mel}}{r_{me}^2}\left[V''(r_{me})-\frac{V'(r_{me})}{r_{me}}\right]+4M''(P_e)\sum_m \frac{r_{mi}}{r_{me}^2}\phi^2(r_{me})\\ &\left[r_{mex}\delta_{xj}+r_{mey}\delta_{yj}+\beta r_{mez}\delta_{zj}-7r_{mej}\frac{r_{mex}^2+r_{mey}^2+\beta r_{mez}^2}{r_{me}^2}\right]\sum_m \frac{r_{mk}}{r_{me}^2}\phi^2(r_{me})\\ &\left[r_{mex}\delta_{xl}+r_{mey}\delta_{yl}+\beta r_{mez}\delta_{zl}-7r_{mel}\frac{r_{mex}^2+r_{mey}^2+\beta r_{mez}^2}{r_{me}^2}\right]\end{aligned} \tag{9-59}$$

据此就可计算 hcp 金属的 5 个独立弹性常数。这样就确定了修正函数所包含的 2 个模型参数（α 和 β）和势函数中的 8 个模型参数（j=−1，0，1，2，3，4，5，6）。

电子密度、嵌入函数和修正函数完全可按照前面的计算方法来确定。hcp 结构金属的晶格常数、结合能、空位形成能和弹性常数等输入参数如表 9-5 所示。

表 9-5　hcp 金属的输入参数

元素	晶格常数		结合能	空位形成能	C_{11}/GPa	C_{12}/GPa	C_{44}/GPa	C_{13}/GPa	C_{33}/GPa
	a/nm	c/nm	E_c/eV	E_{1f}/eV					
Be	0.22856	0.35832	3.32	1.11	282.3	26.7	162.5	14	246
Co	0.24970	0.40690	4.39	1.35	295	159	71	111	335
Hf	0.31946	0.50511	6.44	1.80	181	77	55.7	66	197
Mg	0.32094	0.52105	1.51	0.58	59.3	25.7	16.4	21.4	61.5
Re	0.27600	0.44580	8.03	2.30	616	273	161	206	683
Ru	0.27057	0.42816	6.74	1.85	563	188	181	168	535
Sc	0.33080	0.52670	3.90	1.15	99.3	39.7	27.2	29.4	107
Ti	0.29506	0.46788	4.85	1.50	160	90	46.5	66	181
Y	0.36474	0.57306	4.37	1.25	77.9	28.5	24.31	21.0	76.9
Zr	0.32312	0.51477	6.25	1.70	1.44	74	33.4	67	166

（4）模型参数ϕ_e的确定

MAEAM 模型中使用的电子密度函数是沿用 Johnson 的形式。电子密度函数表达式中的前因子ϕ_e的值原则上可以通过拟合 Clementi 和 Foetti 基于 Hartree−Fock 理论计算的结果来确定。

在实际应用中，ϕ_e的值有两种确定方法：

1）Foiles 的方法。利用合金系统已有的热力学性质来确定合金中各组元素的相对值。

2）Johnson 的方法。采用经验关系式来表示，如定义ϕ_e为能量密度，取$\phi_e = E_c/\Omega$。

Zhan 等在其应用中取$\phi_e = \left[\left(E_c - E_{1f}\right)/\Omega\right]^{3/5}$。胡望宇考虑到能量是电子密度的平方取$\phi_e = \sqrt{E_c}/\Omega$。在将该模型应用于金属元素的计算时，电子密度是以比值形式出现在嵌入能表达式中的，电子密度函数表达式中的前因子几乎被约化掉了，其值的大小不会影响计算结果。但应用于合金的计算时，该参数不能被约化，因而在改进分析型 EAM 模型应用于合金系统时，必须先确定ϕ_e的值。

下面仍采用类似于 Johnson 的方法，通过反复计算和比较，建立电子密度函数表达式中的前因子f_e与元素的物理量之间的关系。

考虑到ϕ_e表示金属元素处于平衡状态时的电子密度，这个量应与元素的电子因素密切相关。它与合金中元素吸引电子的能力有关，即与电负性等物理量相关，电负性大吸引电子的能力强，电子密度大。此外，它也与合金中元素所占体积有关，对于同一电荷数量来说，元素体积大，电子密度小。金属间化合物的有序结构对电子密度的影响我们认为可以通过化学标度来描述。

基于上述考虑，ϕ_e可采用了如下的经验关系

$$\phi_e = c\frac{\phi^* \chi}{r_{1e}^3} \tag{9-60}$$

其中：

1）c为系数，对 AB 类金属间化合物，密堆结构的 Ni、Co、Al 的c取为 1；bcc 金属 Fe，考虑到其非密堆结构的影响，c取为$(0.74/0.68)^3$，对 Ti 元素c取为$\sqrt[3]{E_c}$。对 A_3B 类金属间化合物，由于成分的变化引起电子分布变化，A 元素的ϕ_e计算公式中再乘以一个因子$\sqrt[3]{n_{ws}^{1/3}}$。表 9-6 为一些 AB 型合金和 A_3B 型合金所使用的ϕ_e值。

2）ϕ^*为 Miedema 模型中的参数—功函数，该参量类似于元素的电负性参量，是一个描述元素的电子因素的参量；$\sqrt[3]{n_{ws}^{1/3}}$是 Miedema 模型中定义的 Wigner-Seitz 原胞边界的电子密度，也是一个描述元素的电子因素的参量。

表 9-6　模型参数ϕ_e值

合金	Fe	Co	Ni	Ti	Al
AB 型合金	0.4117	0.3421	0.3660	0.2049	0.2969
A_3B 型合金	0.4960	0.4123	0.4416	—	—

3）χ 为 Pettifor 提出的元素“化学标度”。国际材料科学组织（MSIT）在 1990 年决定采用以 Pettifor 提出的这个参量来确定化学式中元素顺序。这个参量就是所谓的门捷列夫数（M）及相关的“化学标度 χ”。

关于 ϕ_e 的进一步研究，将依据量子力学对原子电子密度的具体计算并考虑元素间合金化效应来计算，这个问题有待日后的深入研究来解决。

2. 金属间化合物的物理性质及点缺陷的计算方法

（1）晶格常数和弹性常数的计算方法

金属间化合物晶格常数的计算：给定金属间化合物的初始晶格常数值，构造出所要研究的晶体结构。使用分子动力学程序对该晶体结构进行弛豫，使其达到平衡状态，此时所得的晶格常数即为本书的计算结果，也是其他物理性质和点缺陷研究的基础。

金属间化合物弹性常数的计算：按 Johnson 的方法对能量展开

$$E = A + \alpha_{ij}A_{ij} + \frac{1}{2}\alpha_{ij}\alpha_{kl}A_{ijkl} + \cdots \tag{9-61}$$

其中：

$$\begin{aligned}
A_{ijkl} &= \Omega C_{ijkl} \\
&= F''(\rho_e)\sum_m \frac{r_{mei}r_{mej}}{r_{me}}\phi'(r_{me})\sum_m \frac{r_{mek}r_{mel}}{r_{me}}\phi'(r_{me}) \\
&\quad + F'(\rho_e)\sum_m \left\{ \begin{aligned} &\frac{r_{mei}r_{mej}r_{mek}r_{mel}}{r_{me}^2}\left[\frac{\partial^2\phi(r_{me})}{\partial r^2} - \frac{1}{r_{me}}\frac{\partial\phi(r_{me})}{\partial r}\right] \\ &+\delta_{jl}\frac{r_{mek}r_{mel}f'(r_{me})}{r_{me}} \end{aligned} \right\} \\
&\quad + \frac{1}{2}\sum_m \left\{ \frac{r_{mei}r_{mej}r_{mek}r_{mel}}{r_{me}^2}\left[V''(r_{me}) - \frac{V'(r_{me})}{r_{me}}\right] + \delta_{jl}\frac{r_{mek}r_{mel}V'(r_{me})}{r_{me}} \right\} \\
&\quad + 4M''(P_e)\sum_m \frac{r_{mei}r_{mej}}{r_{me}}\phi(r_{me})\phi'(r_{me})\sum_m \frac{r_{mek}r_{mel}}{r_{me}}\phi(r_{me})\phi'(r_{me}) \\
&\quad + 2M'(P_e)\sum_m \left\{ \begin{aligned} &\frac{r_{mei}r_{mej}r_{mek}r_{mel}}{r_{me}^2}\left[\phi''(r_{me})\phi(r_{me}) + \phi'(r_{me})\phi'(r_{me}) - \frac{\phi(r_{me})\phi'(r_{me})}{r_{me}}\right] \\ &+\delta_{jl}\frac{r_{mei}r_{mek}\phi(r_{me})\phi'(r_{me})}{r_{me}} \end{aligned} \right\}
\end{aligned} \tag{9-62}$$

式中，Ω 为原胞中原子平均体积。

由已计算出的晶格常数值即可计算出金属间化合物的弹性常数。

（2）金属间化合物的点缺陷及扩散性质的计算

B2 结构金属间化合物 AB 的晶体结构是体心立方为基的有序结构，由一种元素（如 A）占据 bcc 晶胞的体心位置构成的简单立方子晶格（称 α 子晶格），另一元素（如 B）占据晶胞的顶角位置构成的简单立方子晶格（称 β 子晶格），两个子晶格互相穿插构成有序结构。L12 结构金属间化合物 A_3B 的晶体结构是以面心立方为基的有序结构，A 原子占据面心立方中面心的位置，B 占据顶角位置。L10 结构金属间化合物 AB 的晶体结构也是以面心立方为基的有序结构，A、B 原子分别分层占据面心立方中（001）面，沿[001]方向按 ABAB…重复排列而成。

完整晶体结构只在绝对零度理想化学配比成分下出现。在一定温度下，晶体中出现的点缺陷称为热缺陷：为容纳合金与理想化学配比成分的偏离出现的点缺陷称为结构缺陷。在这种有序结构中可能出现四种结构点缺陷：A 原子占据 β 子晶格形成 A 反位置缺陷，B 原子占据 α 子晶格形成 B 反位置缺陷，α 子晶格中形成 A 空位和 β 子晶格形成 B 空位。

模拟过程中得到的各种点缺陷能量由下述公式给出。

对过渡元素（TM）和 Al 原子的单空位

$$
\begin{aligned}
E_v(\mathrm{TM}) &= E_{rel}(n-1) - E_{per}(n) + E_{coh}(\mathrm{TM}) \\
E_v(\mathrm{Al}) &= E_{rel}(n-1) - E_{per}(n) + E_{coh}(\mathrm{Al})
\end{aligned}
\tag{9-63}
$$

对过渡元素（TM）和 Al 原子的双空位

$$
\begin{aligned}
E_{2v}(\mathrm{TM}-\mathrm{TM}) &= E_{rel}(n-2) - E_{per}(n) + 2E_{coh}(\mathrm{TM}) \\
E_{2v}(\mathrm{TM}-\mathrm{Al}) &= E_{rel}(n-2) - E_{per}(n) + E_{coh}(\mathrm{TM}) + E_{coh}(\mathrm{Al}) \\
E_{2v}(\mathrm{Al}-\mathrm{Al}) &= E_{rel}(n-2) - E_{per}(n) + 2E_{coh}(\mathrm{Al})
\end{aligned}
\tag{9-64}
$$

对过渡元素（TM）和 Al 原子的反位置缺陷

$$
\begin{aligned}
E_{anti}(\mathrm{TM}\to\mathrm{Al}) &= E_{rel}(n) - E_{per}(n) + E_{coh}(\mathrm{Al}) - E_{coh}(\mathrm{TM}) \\
E_{anti}(\mathrm{Al}\to\mathrm{TM}) &= E_{rel}(n-1) - E_{per}(n) + E_{coh}(\mathrm{TM}) - E_{coh}(\mathrm{Al})
\end{aligned}
\tag{9-65}
$$

式中，E_v 表示单空位的形成能，E_{2v} 表示双空位的形成能，E_{anti} 表示反位置缺陷的形成能，$E_{per}(n)$表示 n 个原子组成的完整晶格的总能量，$E_{rel}(n\text{-}1)$表示引入一个点缺陷后余下的（n-1）个原子组成晶格的总能量，E_{coh}(TM)和 E_{coh}(Al)表示单个 TM 原子和 Al 原子的结合能。

金属间化合物中原子扩散的研究：根据扩散空位机制，原子通过空位而迁移，邻近空位的原子向空位迁移时必须克服周围原子相互作用所构势垒的束缚，通过计算发现势垒的最高点位于原子迁移路途的中点，迁移原子处于此位置时的能量与处于平衡位置的能量差即为空位迁移能 $E_{1m} = E_{1f} - E_0$。原子的扩散激活能 Q_{1v}，

$Q_{1v} = E_{1f} + E_{1m}$。

分别计算了原子扩散过程弛豫和未弛豫的能量。未弛豫的原子扩散研究是分别计算扩散路径各点的能量及扩散过程的迁移能和激活能，没有考虑分子动力学弛豫。弛豫的原子扩散研究是把原子放在扩散路径的鞍点上，作分子动力学弛豫，原子将在鞍点作弛豫并逐渐向空位移动，最后到达格点位置，在弛豫过程中对各点的能量取样即可计算出原子扩散的迁移能和激活能。

9.5 Finnis-Sinclair（F-S）多体势

F-S 多体势最初是 1984 年由 Finnis 和 Sinclair 基于金属能带的紧束缚理论发展来的，在数学上等同于 EAM 的势函数。Finnis−Sinclair（F−S）多体势和 EAM 模型较成功地描述了金属及合金中的原子间的相互作用势。EAM 和 F−S 方法的主要不同在于对势函数的多体部分的来源和解释不同。本节重点阐述 F−S 多体势的发展、构建及应用。

9.5.1 F−S 多体势的初始模型

N 个原子系统的总能量可以表示为

$$\begin{aligned} E &= E_P + E_N \\ &= \frac{1}{2}\sum_{ij} V(r_{ij}) - A\sum_i f(\rho_i) \end{aligned} \tag{9-66}$$

式中，E_P 是中心对势项，代表核−核间的排斥作用，E_N 代表多体相互作用，A 是大于零的常数，通常取为 1。

$f(\rho)$的形式如下：

$$f(\rho_i) = \sqrt{\rho_i} \tag{9-67}$$

其中，$\rho_i = \sum_{j \neq i} \phi(r_{ij})$。

这里，$r_{ij} = |r_i - r_j|$，式中 r_i 是原子的空间坐标；在紧束缚理论中函数$\phi(r)$可以解释为重积分的平方和。

为了对粒子体系模拟计算的方便，电子电荷的密度函数ϕ和满壳电荷的排斥势函数 V 的形式必须简单而又能表征其特性，函数$\phi(r)$采用抛物线的形式

$$\phi(r) = \begin{cases} (r-d)^2 & r \leqslant d \\ 0 & > d \end{cases} \tag{9-68}$$

式中，d 为截断距离，位于第二和第三近邻之间，即 $a < d < \sqrt{2}a$，a 是晶格常数，对于金属 Cr 和 Fe，式（9-68）被修正为

$$\phi(r) = (r-d)^2 + \frac{\beta}{d}(r-d)^3 \tag{9-69}$$

式中，参数的选择原则是使$\phi(r)$在第一近邻距离内达到一个最大值，即要求：如果d=a，则取$\beta < 4.975$；如果$d = 2$，则取$\beta < 4.975$。

对势函数V采用四次多项式的形式

$$V(r)=\begin{cases}(r-c)^2\left(c_0+c_1r+c_2r^2\right) & r\leqslant c\\ 0 & >c\end{cases} \tag{9-70}$$

式中，c也是截断距离，同d一样也定在第二和第三近邻之间。而c_0、c_1、c_2则必须通过拟合具体材料的实验数据获得。Finnis 等人拟合得到的几种过渡金属的 F-S 多体势参数见表 9-7。

表 9-7　Finnis 等人拟合得到的 V、Ni、Ta、Cr、Mo、W、Fe 的 F-S 多体势参数

元素	d / nm	β	c	c_0	c_1	c_2
V	0.3692767	0	3.80	−0.8816318	1.4907756	−0.3976370
Ni	0.3915354	0	4.20	−1.5640104	2.0055779	−0.4663764
Ta	0.4076980	0	4.20	1.2157373	0.0271471	−0.121735
Cr	0.3915720	1.8	2.90	29.1429813	−23.3975027	4.7578297
Mo	0.4114825	0	3.25	43.4475218	−31.9332978	6.0804249
W	0.4400224	0	3.25	47.1346499	−33.7665655	6.2541999
Fe	0.3699579	1.8	3.4	1.2110601	−0.7510840	0.1380773

9.5.2　改进的 F-S 多体势模型

Ackland 等人发现式（9-65）和式（9-67）所给出的对势和电子密度函数的形式有缺点，主要表现在：在运用分子动力学模拟计算时，原子会像受到高压一样最终被挤到一块，势函数如何在短距离内表现为吸引作用。为克服这个缺点，他们对 F-S 多体势函数做了一些改进，即在对势的表达式（9-67）中又增加了一指数项，目的是在 F-S 多体势的对势排斥部分中通过加入一个芯项来增加短程斥力。具体如下式

$$V(r)=V_{FS}(r)+B\left(b_0-r\right)\mathrm{e}^{-\alpha r} \tag{9-71}$$

式中，b_0是最近邻距离，通过拟合区域$0.4 < r/b_0 < 1.0$来确定参数B和α。Ackland 等人还在此基础上拟合了几种金属修正的多体势参数，见表 9-8。

表 9-8　金属 V、Nb、Ta、Mo、W 的修正的对势参数

元素	B/(E_v/nm)	α/nm	r_1/nm
V	23000	5	0.26320
Nb	48000	8	0.28585
Ta	91000	10.5	0.28629
Mo	1223000	39	0.27255
W	90300	12	0.27411

9.5.3 立方样条函数的 F-S 多体势模型

1987 年，Ackland 等人提出了更加方便的三次样条函数形式的 F-S 多体势，如下式所示：

$$
\begin{aligned}
V(r) &= \sum_{k=1}^{6} \alpha_k \left(r_k - r\right)^3 H\left(r_k - r\right) \\
\phi(r) &= \sum_{k=1}^{2} A_k \left(R_k - r\right)^3 H\left(R_k - r\right)
\end{aligned}
\tag{9-72}
$$

式中，$H(x)$是阶跃函数，当 $x > 0$ 时，$H(x) = 1$；否则，$H(x) = 0$。r_k 和 R_k 是函数的节点值，选择的原则是 $r_1 > r_2 > r_3 > r_4 > r_5 > r_6$，$R_1 > R_2$。

Ackland 等人根据上述模型，拟合得到了 Cu、Ag、Au、Ni 四种纯金属的 F-S 多体势参数，如表 9-9 所示。

表 9-9 Ackland 等人拟合得到了 Cu、Ag、Au、Ni 四种纯金属的 F-S 多体势参数

参数	Cu	Ag	Au	Ni
α_1	29.059214	20.368404	29.059066	29.057085
α_2	−140.05681	−102.36075	−153.14779	−76.04625
α_3	130.07331	94.31277	148.17881	48.08920
α_4	−17.48135	−6.220051	−22.20508	−25.96604
α_5	31.82546	31.08088	72.71465	79.15121
α_6	71.58749	175.56047	199.26269	—
r_1	1.2247449	1.2247449	1.2247449	1.2247449
r_2	1.1547054	1.1547054	1.1547054	1.1547054
r_3	1.1180065	1.1180065	1.1180065	1.1180065
r_4	1.0000000	1.0000000	1.0000000	1.0000000
r_5	0.8660254	0.8660254	0.8660254	0.8660254
r_6	0.7071068	0.7071068	0.7071068	0.7071068
R_1	1.2247449	1.2247449	1.1180065	1.1180065
R_2	1.0000000	1.0000000	1.0000000	1.0000000
A_1	9.806694	1.458761	21.930125	60.537985
A_2	16.774638	42.946555	284.99631	−80.102414

9.5.4 合金中的 F-S 多体势

最初建立的 F-S 多体势主要应用于纯金属，直到 1987 年 Ackland 等人首次将 F-S 多体势应用于二元合金系统，之后很多研究人员在该模型下拟合了适应于不同的二元合金系统的原子间相互作用势。1999 年 Landa 等人又将 F-S 多体势首次

推广到了 Pb−Bi−Ni 三元合金系统。除此之外，由于多元合金的复杂性，目前所建立的原子间互作用势仍然只适用于纯金属、金属间化合物及二元合金。

1. 二元合金的 F−S 多体势

二元合金的 F−S 多体势仍采用式（9-63）～式（9-65）所示的形式，其中对势及多体势函数可用下面 6 个函数表示，即：V_{AA}、V_{BB}、V_{AB}、ϕ_{AA}、ϕ_{BB}、ϕ_{AB}，其中 A、B 代表原子的种类。同时假定 V_{AA}、V_{BB}、ϕ_{AA}、ϕ_{BB} 与浓度有关，因此可以认为 V_{AA}、V_{BB}、ϕ_{AA}、ϕ_{BB} 和纯金属的参数相同。为计算方便，函数 ϕ_{AB} 取 ϕ_{AA} 和 ϕ_{BB} 的几何平均，因此只有对势项 V_{AB} 需要通过拟合合金已知的物理参数而得到。

为保持和纯金属的 F−S 多体势函数形式一致，V_{AB} 也采用立方函数形式，即构成上述模型的各个函数可写为

$$
\begin{aligned}
V_{AA}(R_{ij}) &= \sum_{k=1}^{6} \alpha_k^{AA} H(r_k^{AA} - R_{ij})(r_k^{AA} - R_{ij})^3 \\
V_{AB}(R_{ij}) &= \sum_{k=1}^{5} \alpha_k^{AB} H(r_k^{AB} - R_{ij})(r_k^{AB} - R_{ij})^3 \\
\phi_{AA}(R_{ij}) &= \sum_{k=1}^{2} \alpha_k^{AA} H(R_k^{AA} - R_{ij})(R_k^{AA} - R_{ij})^3 \\
\phi_{AB}(R_{ij}) &= \sqrt{\phi_{AA}(R_{ij})\phi_{BB}(R_{ij})}
\end{aligned}
\tag{9-73}
$$

其中 $H(x)$ 仍是阶跃函数。V_{AA}、ϕ_{AA} 和 V_{BB}、ϕ_{BB} 具有相同的形式。

大多数金属间化合物都具有很高的高温比强度和比模量，同时还具有抗高温、抗氧化、低密度等优点，如：Ni_3Al、NiAl 等都是新型的金属基高温材料，它们不仅具有上述优良性质，而且还是具有广泛应用前景的轻质高温材料，因而越来越受到人们的关注。Ackland 等人在 1987 年构建了 Ni_3Al 的 F−S 多体势，并拟合了它的多体势参数，见表 9-10。

表 9-10　Ni_3Al 的 F−S 多体势参数

参数	Ni	Al	Ni−Al
α_1	32.9100	32.2588	−0.64692088
α_2	−51.3477	−74.5610	1.13926928
α_3	39.2079	140.0125	−0.66551061
α_4	−88.2142	−107.5789	1.46802193
α_5	133.1403	29.4447	3.0000000
α_6	200.0000	101.5500	—
r_1	1.225	1.225	4.35174
r_2	1.202	1.180	4.24473
r_3	1.100	1.090	3.88803
r_4	0.900	1.050	2.96061

（续）

参数	Ni	Al	Ni−Al
r_5	0.820	0.900	2.52223
r_6	0.69	0.707	—
R_1	1.225	1.200	—
R_2	0.930	0.93000	—
A_1	28.7717	19.2224	—
A_2	−94.0309	4.9695	—

2. 三元合金的 F-S 多体势

1999 年，Landa 等人建立了 Pb−Bi−Ni 三元合金系统的 F−S 多体势，其形式如下

$$E=\sum_{i=1}^{N}\left[\frac{1}{2}\sum_{j=1}^{N}V_{S_iS_j}(R_{ij})-\sqrt{\rho_{S_i}}\right] \tag{9-74}$$

式中，$\rho_{S_i}=\sum_{j=1}N\phi_{S_iS_j}(R_{ij})$。

对于三元合金，V 和 Φ 分别需要 6 种不同的函数 V_{AA}、V_{BB}、V_{CC}、V_{AB}、V_{BC}、V_{AC}、Φ_{AA}、Φ_{BB}、Φ_{CC}、Φ_{AB}、Φ_{BC}、Φ_{AC} 其中 V_{AA}、V_{BB}、V_{CC}、Φ_{AA}、Φ_{BB}、Φ_{CC} 采用与纯金属对应的形式。函数 Φ_{AB} 取 Φ_{AA} 与 Φ_{BB} 的几何平均，Φ_{BC}、Φ_{AC} 与 Φ_{AB} 的形式一样。Pb−Bi−Ni 三元合金系统的 F−S 多体势参数见表 9-11。

表 9-11　Pb−Bi−Ni 三元合金系统的 F−S 多体势参数

参数	Ni	Bi	Pb	Pb−Bi	Pb−Ni	Ni−Bi
α_1	32.9100	0.0000	7.51545	0.023237	−0.020654	0.078778
α_2	−51.3477	−58.9819	−16.1876	−0.062135	0.121422	−0.163839
α_3	39.2079	123.1985	10.8639	0.165322	4.155184	0.956508
α_4	−88.2142	−94.7936	−2.7654	—	—	—
α_5	133.1403	96.2540	29.0132	—	—	—
α_6	200.0000	30.7920	00000	—	—	—
r_1	1.225	1.250	1.250	6.1248	5.5900	5.280
r_2	1.202	1.200	1.200	5.8812	4.5700	4.884
r_3	1.100	1.150	1.110	4.9416	3.2700	3.498
r_4	0.900	1.050	0.900	—	—	—
r_5	0.820	0.900	0.880	—	—	—
r_6	0.707	0.866	0.700	—	—	—
R_1	1.225	1.250	1.250	—	—	—
R_2	0.930	1.050	1.250	—	—	—
A_1	28.7717	7.51283	1.82546	—	—	—
A_2	−94.0309	−9.79580	40.32900	—	—	—

9.6　晶格反演方法（LIM）

如前所述，原子间势函数的获得普遍采用的是多参数拟合的办法，这也是在凝聚态物理、材料物理中处理复杂计算问题的常规办法。原子间势模型中经常包含很多待定参数，这些参数通过拟合已知结构的力学、热力学数据来确定，拟合程度的好坏在一定程度上体现势模型的合理性。如对势模型，Cauchy 关系是径向二体势的严格结果，因此除离子晶体或惰性气体元素晶体对势模型外，一般不能拟合出其他材料的弹性常数，这反映出对势模型的局限性。多参数拟合存在一个严重的问题就是拟合的结果存在一定的不确定性，因为从势模型计算材料的力学、热力学量大多是复杂的非线性关系式，非线性方程组的拟合本身就包含有内在的不稳定性和不唯一性。实际的拟合计算中往往根据物理的直觉给拟合参数限定范围，并给定初值，而拟合结果对约束与初值的改变都很敏感，这是非线性拟合的内禀性质，对拟合结果的合理性与可信度都有一定影响。如果所涉及的是人工设计的新材料，没有可得到的实验数据，参数拟合的方法则不能用。晶格反演方法（lattice inversion method 或缩写为 LIM）——从系统的总能曲线（势能作为体积或晶格常数的函数）直接反演原子间相互作用势函数的方法——为势函数的计算提供了一条有效的新途径，Chen−Möbius 定理将数论中的 Möbius 变换推广到应用物理中逆问题的处理，特别是格子体系反演问题的求解，使得周期结构中对势的反演问题得到了最完美的解决。

9.6.1　Carlsson-Gelatt-Ehrenreich（CGE）公式

我们先从 CGE 公式的最原始推导开始：考虑一个由单一原子组成的晶体结构，在对势模型下，晶体的总势（每个原子的平均结合能）按径向二体势的求和可以表示成

$$E(x)=\frac{1}{2}\sum_{|\boldsymbol{R}_i|\neq 0}V(|\boldsymbol{R}_i|)=\frac{1}{2}\sum_{n=1}^{\infty}w(n)\,V\left(q(n)x\right) \tag{9-75}$$

$|\boldsymbol{R}_i|$ 表示晶体中的原子与坐标原点的参考原子之间的位矢，x 表示最近邻原子间距，$w(n)$表示半径为 $q(n)x$ 壳层中的原子数目，并且 $q(1)=1<q(2)<q(3)<\cdots$ 按编号排序。定义加权的标度变换算符 T_p，对任意的函数 $f(x)$

$$T_p f(x)=w(p)f\left(q(p)x\right) \tag{9-76}$$

因为 $q(1)=1$，所以 $T_1^{-1}=\dfrac{1}{w(1)}$。将方程（9-75）用算符表示，则有

$$E(x)=\frac{1}{2}\sum_{n=1}^{\infty}T_nV(x)=\left[T_1\left(1+\sum_{n=1}^{\infty}T_1^{-1}T_n\right)\right]V(x) \tag{9-77}$$

反过来，即得到下面的反演公式

$$
\begin{aligned}
V(x) &= 2\left(1+\sum_{p=2}^{\infty}T_p\right)^{-1}T_1^{-1}E(x) \\
&= 2\left(1-\sum_{p=2}^{\infty}T_1^{-1}T_p+\sum_{p,q=2}^{\infty}T_1^{-1}T_pT_1^{-1}T_q-\cdots\right)T_1^{-1}E(x) \\
&= \frac{2}{w(1)}E(x)-\sum_{p=2}^{\infty}\frac{2w(p)}{w(1)^2}E\big(w(p)x\big)+\sum_{p,q=2}^{\infty}\frac{2w(p)w(q)}{w(1)^3}E\big(w(p)w(p)x\big)-\cdots
\end{aligned}
\tag{9-78}
$$

在晶体中，$E(x)$一般为指数衰减函数，因此可以保证展开式的收敛性，而在截断近似下，式（9-78）变为一个有限的级数。值得一提的是，利用反演公式不仅可直接从第一原理的总势计算反演对势函数，而且势函数包含了晶体在非平衡态下的性质及长程力行为，克服了势参数拟合的经验做法中仅仅包含平衡晶格性质的缺陷。

9.6.2 LIM 与异种原子间的对势反演

晶格反演方法同样可用来反演异种原子间的相互作用二体势函数。设两种原子组成的晶体，总势

$$
E = E_{AA}+E_{BB}+E_{AB}=\sum V_{AA}+\sum V_{BB}+\sum V_{AB} \tag{9-79}
$$

其中，相同原子间的相互作用势 V_{AA} 和 V_{BB} 可以分别计算由 A、B 两种原子组成的单原子晶体总势反演得到，因此异种原子间的二体势函数$\phi_{A\text{-}B}$即可从偏结合能 E_{AB} 反演。这里暗含了反演的对势具有移植性的假设，因为反演强调的是一种普适的方法，先不涉及势函数的合理性问题。类似的办法可以推广到任意多元化合物晶体中的对势反演。设 N 元化合物晶体，涉及 $N(N+1)/2$ 种不同的对势函数，因此至少要计算 $N(N+1)/2$ 条不同的包含这种化合物的元素的不同晶体结构的总势曲线。如二元金属间化合物 Ni_3Al，涉及到三种不同的对势函数（即 V_{NiNi}、V_{AlAl} 和 V_{NiAl}）。通过计算三种不同结构的总势曲线（如选用 fcc−Ni，fcc−Al 和 L12−Ni_3Al），即可求出所涉及的三种对势函数。具体计算时一般先用模型函数拟合总势曲线和反演的对势曲线，如在反演ϕ_{NiAl}时，实际上先进行了若干次曲线拟合（如 $E_{\text{fcc−Ni}}$、$E_{\text{fcc−Al}}$、$E_{\text{L12−Ni3Al}}$、V_{NiNi}、V_{AlAl} 等）。对于多元化合物晶体，过多的中间拟合过程会带来一定的累积误差，因此可以先对中间过程作严格的算子运算，以消除中间拟合误差。

设某个问题中要计算 N 种不同的对势函数，通过计算 N 条不同的总势曲线得到如下的关系式

$$
\begin{aligned}
E_1 &= S_{11}V_1+S_{12}V_2+\cdots+S_{1n}V_n \\
E_2 &= S_{21}V_1+S_{22}V_2+\cdots+S_{2n}V_n \\
&\vdots \\
E_n &= S_{n1}V_1+S_{n2}V_2+\cdots+S_{nn}V_n
\end{aligned}
\tag{9-80}
$$

每一个算子实际上都是格子体系中的不同形式的求和。令

$$E=\begin{bmatrix} E_1 \\ E_2 \\ \cdots \\ E_n \end{bmatrix} \quad V=\begin{bmatrix} V_1 \\ V_2 \\ \cdots \\ V_n \end{bmatrix} \quad S=\begin{bmatrix} S_{11} & S_{12} & \cdots & S_{1n} \\ S_{21} & S_{22} & \cdots & S_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ S_{n1} & S_{n2} & \cdots & S_{nn} \end{bmatrix}$$

则式（9-79）及其对势反演可以简写为：$E=SV_i$ 及 $V=S^{-1}E$。

因为求和算子彼此对易，因此矩阵 S 及其求逆的运算与代数矩阵的逆运算完全相同，其中涉及到算子之间的加减乘除运算（除即乘算子的逆），而每一个算子 S_{ij} 的逆都可以从 Chen-Möbius 定理求得。

9.6.3 LIM 与原子间的多体势反演

如上所述，LIM 为从第一原理计算到原子间（包括相同原子与不同原子间）的对势反演提供了简洁有效的方法，但如何用 LIM 反演原子间的多体势函数仍然是一个有待解决的问题。例如，如果晶体总势求和中包含三体势，即

$$\begin{aligned} E_c(r) &= \frac{1}{2}\sum_i V_2(r_i) + \frac{1}{6}\sum_{i,j} V_3(r_i, r_j, r_{ij}) \\ &= \sum_{n=1}^{\infty} w(n) V_2\left(r(n)r\right) + \frac{1}{6}\sum_{i,j} V_3(r_i, r_j, r_{ij}) \end{aligned} \tag{9-81}$$

其中 V_2 和 V_3 分别为对势和三体势函数，显然我们不可能从一个方程反演出两个未知函数。即使对确定的原子系统，可以从原子组成的不同晶体结构的总势计算得到如式（9-81）求和形式的方程，三体势函数仍然不能反演，原因在于三体势是包含有三个独立自变量的函数。

将式（9-81）写成算符形式

$$E_c(r) = R_2 V_2(r) + R_3 V_3(r_1, r_2, r_3) \tag{9-82}$$

由 LIM，我们得到

$$V_2(r) = T_\mu E_c(r) + T_\mu R_3 E_c(r) \tag{9-83}$$

基于上式的一种半经验的办法是将三体部分采用含经验参数的经验模型势表示，然后由第一原理的总势计算反演出对势函数，完全第一原理的做法可先对三体的模型势进行修改，三体势参数也由第一原理的计算得出，然后做第一原理的对势反演。

在多体势分析中，多体势函数都表示成自变量的分离变量形式，LIM 也可以直接反演多体势中的未知函数。以 s-价电子系统为例

$$E_c(r) = E_{ion-ion}(r) + \sum_{n=1}^{\infty} w(n) h^2\left(r(n)r\right) + \frac{1}{6}\sum_{ij} C^{(3)} h(r_i) h(r_j) h(r_{ij}) + \cdots \tag{9-84}$$

其中，未知函数域 $h(r)$ 可从如下反演方程做自洽计算得到，即

$$h(r) = -\sqrt{T_\mu\left[E_c(r) - E_{ion-ion}(r) - \frac{1}{6}\sum_{ij} C^{(3)} h(r_i) h(r_j) h(r_{ij}) - \cdots\right]} \qquad (9\text{-}85)$$

但对于具体的系统，首先要以 Slater−Koster 关系为依据，对 $C^{(n)}$进行简化，确定与原子位形无关的多体势参数。

9.6.4 陈氏三维晶格反演（Chen−Möbius）

晶体的结合能一般可表示为原子间势函数的求和

$$E(x) = \frac{1}{2}\sum_{i \neq j} V(R_{ij}) + \frac{1}{6}\sum_{\substack{k \neq i,j \\ j \neq k}} V_2(R_{ijk}) + \sum V_3(R_{ijkl}) + \cdots \qquad (9\text{-}86)$$

式中右边第一项为二体势项，即对势项，其后是三体势项、四体势项等等。在很多情况下，对势项对结合能的贡献占主导地位。特别是对金属而言，由于其结合力主要来自共有化电子气作用，用与方向无关的二体势来表示原子间相互作用，能够满足很多的应用需求。

从前面几节可以看出，经验的原子间相互作用二体势有很多形式，并有较大的不确定性，不同的理论依据、不同的近似方法，会得到不同的结果。因此，寻找理论依据充分、计算方法严谨的获取二体势方法是非常必要的。

陈难先等通过一系列的第一原理能带计算，得到各种材料详尽的结合能曲线，运用以虚拟结构和数论方法为基础的晶格反演理论，依据晶体结合能曲线得出多种晶体结构的反演系数，建立起具有一定规模的有效的原子间相互作用势库，成为在原子尺度上进行材料模拟设计的有力工具。从而可以由材料结合能曲线得到同种及异种原子间相互作用对势并运用于 EAM 模型，用于各种复杂体系的计算。

1. 陈氏三维晶格反演

Maddox 在《Nature》上的文章“Möbius and problems of inversion”，对 Chen 的 1990 年的工作“Modified Möbius Inverse Formula and Its Applications in Physics”给予了高度的评价，并提出了一个具有挑战性的问题：“…这将导致用 Chen 的思想，通过 Möbius 框架解决一些未曾解决的问题，…，一个更迫切的问题是，能否把这种方法（注：通过一维晶格的总势能反演原子间相互作用势的一维晶格反演方法）推广到不仅仅是一维的情况。” Chen 等人接受了这个挑战性的问题，并给出了三维晶格反演的严格的解析结果，这就是陈氏三维晶格反演。在凝聚态物理领域，它得到了广泛的关注和应用，M .Z. Bazant 的报告“Möbius series in version in number theory and physics”对“Chen-Möbius formula”和陈氏三维晶格反演在凝聚态物理领域的贡献进行了分析和阐述。

从第一原理结合能曲线，运用三维晶格反演方法，可以严格地导出原子间相互作用对势。陈氏晶格反演的理论要点如下。

对于同种原子构成的晶体，设从第一原理算出的晶体结合能函数可表示为

$$E(x)=\frac{1}{2}\sum_{n=1}^{\infty}r_0(n)V\left(b_0(n)x\right) \tag{9-87}$$

式中，x 表示最近邻原子间距，$r_0(n)$是第 n 级近邻的配位数；$b_0(n)$是以最近邻距离为单位的、第 n 级近邻到参考原子的距离；$V(x)$为对势函数。为了理论上处理的方便，通过低 $b_0(n)$的自乘得到$\{b(n)\}$，$\{b(n)\}$可认为构成乘法半群，这里多出的虚格点相应的虚配位数均为零。这时有

$$E(x)=\frac{1}{2}\sum_{n=1}^{\infty}r(n)V\left(b(n)x\right) \tag{9-88}$$

这里，$r(n)=\begin{cases}r_0\left(b_0^{-1}\left[b(n)\right]\right), & b(n)\in\left\{b_0(n)\right\}\\ 0, & b(n)\notin\left\{b_0(n)\right\}\end{cases}$，由此可反演出的原子间相互作用势普遍公式

$$V(x)=2\sum_{n=1}^{\infty}I(n)E\left(b(n)x\right) \tag{9-89}$$

式中，$I(n)$可通过下式求出

$$\sum_{b(d)|b(n)}I(d)r\left(b^{-1}\left[\frac{b(n)}{b(d)}\right]\right)=\delta_{n1} \tag{9-90}$$

这里求和号下 $b(d)|b(n)$表示对所有这样的 $b(d)$求和：即 $b(d)\in\left\{b(n)\right\}$，且可以在$\left\{b(n)\right\}$中找到一个$\left\{b(k)\right\}$，使得 $b(d)b(k)=\left\{b(n)\right\}$。

可以证明，式（9-89）是式（9-87）的解

$$\begin{aligned}2\sum_{n=1}^{\infty}I(n)E\left(b(n)x\right)&=2\sum_{k=1}^{\infty}\left\{\sum_{b(n)|b(k)}I(n)r\left(b^{-1}\left[\frac{b(k)}{b(n)}\right]\right)\right\}\phi\left(b(k)x\right)\\&=\sum_{k=1}^{\infty}\delta_{kl}\phi\left(b(k)x\right)\\&=\phi\left(b(1)x\right)\\&=\phi(x)\end{aligned} \tag{9-91}$$

2. 陈氏晶格反演的应用

基于第一性原理计算和陈氏晶格反演获得的原子相互作用对势，不依赖于经验的形式和方法，直接由金属和金属间化合物的第一性原理自洽能带计算和严格的晶格反演过程得出，其准确性仅仅取决于第一性原理计算的准确度和对势近似的可靠性，因此它在材料的物理性质的理论计算中得到了广泛的应用。

基于第一性原理材料结合能的计算结果，利用陈氏三维晶格反演，得到的原子间相互作用势，可以应用于计算材料的静力学和动力学特性，如：利用原子势，可以计算 Ni_3Al 的声子谱和弹性常数，金刚石类型材料的声子谱，Mo、Cr 和 Cu

等过渡族金属的声子谱，Li、Na、K、Rb 等碱金属的声子谱，以及 Fe_3A1 声子谱和振动熵等，并且计算结果和实验结果基本一致。

Ni_3A1 是航空工业十分关注的一种金属间化合物，其第三元素的添加和材料的改性得到了广泛的研究，有关其择优代位问题已有大量实验，其中块状材料择优代位的实验约 30 多种。应用陈氏三维晶格反演得到的原子间相互作用势，对这一问题进行了理论上系统的研究和预测。通过对周期表中元素的替位行为进行广泛的预测和对比，这一方法的计算结果表明，除 Co、Cu、Cr、Fe、Mn、Pd、Pt、Ir、Os、Tc、Ru、Rh、Re 代 Ni 位外，其他元素都表现出代 Al 位的趋势。与已有的 30 多种择优代位实验结果相比较，除 Sc 以外，其余计算结果都与实验一致。对于一些没有实验数据的元素，如 Hg、Bi、Eu、Gd、Dy，则通过计算预测了它们的代位趋向。同时还预言了晶界择优代位和表面择优代位趋势。

参 考 文 献

[1] 程守洙, 江之水. 普通物理学[M]. 北京: 高等教育出版社, 1998.

[2] 向一和. 大学物理导论[M]. 北京: 清华大学出版社, 1997.

[3] 野村昭一郎. 量子力学入门[M]. 李彬，黄东律, 等译. 北京: 高等教育出版社, 1985.

[4] B. N. 瑞德泥克. 量子力学史话[M]. 黄宏荃，彭竟灏, 译. 北京: 科学出版社, 1979.

[5] 张怿慈. 量子力学简明教程[M]. 北京: 人民教育出版社, 1979.

[6] 张三慧. 大学物理—量子物理[M]. 北京: 清华大学出版社, 2000.

[7] 汪志诚. 热力学 • 统计物理[M]. 北京: 高等教育出版社, 2004.

[8] 犬石嘉雄, 滨川圭弘, 白藤纯嗣. 半导体物理[M]. 张志杰, 郗小林, 雷京贵, 译.北京: 科学出版社, 1986.

[9] 胡承正. 统计物理学[M]. 武汉: 武汉大学出版社, 2004.

[10] 苏汝铿. 统计物理学[M]. 北京: 高等教育出版社, 2004.

[11] 俞雪珍, 李法和, 陈光旨. 理论物理概论—量子力学[M]. 北京: 高等教育出版社, 1988.

[12] 沈光球, 陶家洵, 徐功骅. 现代化学基础[M]. 北京: 清华大学出版社, 1999.

[13] 汪复兴. 金属物理[M]. 北京: 机械工业出版社, 1981.

[14] 刘国勋. 金属学原理[M]. 北京: 冶金工业出版社, 1979.

[15] 胡赓祥, 钱苗根. 金属学[M]. 上海: 上海科学技术出版社, 1980.

[16] 王润. 金属材料物理性能[M]. 北京: 冶金工业出版社, 1993.

[17] James P Schaffer. 工程材料科学与设计[M]. 余永宁, 强文江, 等译. 北京: 机械工业出版社, 2003.

[18] 虞丽生. 半导体异质结物理[M]. 北京: 科学出版社, 1990.

[19] 王国梅, 万发荣. 材料物理[M]. 武汉: 武汉理工大学出版社, 2004.

[20] 田莳. 材料物理性能[M]. 北京: 北京航天航空大学出版社, 2004.

[21] 陈金富. 固体物理学[M]. 北京: 高等教育出版社, 1986.

[22] 方俊鑫, 陆栋. 固体物理学[M]. 上海: 上海科学技术出版社, 2001.

[23] 顾秉林, 王喜坤. 固体物理学[M]. 北京: 清华大学出版社, 1989.

[24] 关振铎, 张中太, 焦金生. 无机材料物理性能[M]. 北京: 清华大学出版社, 1992.

[25] 邱成军, 王元化, 王义杰. 材料物理性能[M]. 哈尔滨: 哈尔滨工业大学出版社, 2003.

[26] 黄昆, 韩汝琦, 固体物理学[M]. 北京: 高等教育出版社, 1988.

[27] 曾谨言. 量子力学[M]. 北京: 科学出版社, 1990.

[28] 谢希德, 陆栋. 固体能带理论[M]. 上海: 复旦大学出版社, 1998.

[29] C. 基泰尔. 固体物理导论 •8 •版[M]. 项金钟, 吴兴惠, 译. 北京: 化学工业出版社, 2005.

[30] 张跃, 谷景华, 尚家香, 等. 计算材料学基础[M]. 北京: 北京航空航天大学出版社, 2007.

[31] 龙永强. Cu-Ni-Si 合金微结构及相变的计算与模拟[D]. 上海: 上海交通大学, 2011.

[32] 潘洪哲. 氮化硅材料的第一性原理研究[D]. 成都: 四川师范大学, 2006.
[33] 黄美纯. 密度泛函理论若干进展[J]. 物理学进展, 2000, 20(3): 199-219.
[34] 姚强, 邢辉, 孟丽君, 等. TiB_2 和 TiB 弹性性质的理论计算[J]. 中国有色金属学报, 2007, 17(8): 1277-1281.
[35] J C Slater. An Augmented Plane Wave Method for the Periodic Potential Problem [J]. Phys. Rev., 1953, 93(3): 603-608.
[36] P Hohenberg, W. Kohn. Inhomogencous electron gas [J]. Phys. Rev. B, 1964, 136: 864-871.
[37] W Kohn, L J Sham. Self−consistent equations including exchange and correlation effects[J], Phys. Rev. A, 1965, 140(4A): 1133-1138.
[38] J Callaway, N H March. Density functional methods: theory and applications[J]. Solid State Physics, 1984, 38: 136-223.
[39] J P Perdew. Accurate density functional for the energy real−space cut off of the gradient expansion for the exchange hole [J]. Phys. Rev. Lett., 1985, 55(16): 1665-1668.
[40] J P Perdew, Y Wang. Accurate and simple density fuctional for the electronic exchange energy: generalized gradient approximation [J]. Phys. Rev. B, 1986, 33(12): 8800-8802.
[41] J P Perdew, K Burke, M Ernzerhof. Generalized gradient approximation made simple[J]. Phys. Rev. Lett., 1996, 77(18): 3865-3868.
[42] M W Finnis, J E Sinclair. A simple empirical N−body potential for transition metals[J]. Phil. Mag., 1984, 50(1): 45-55.
[43] G J Ackland, R Thctford. An improved N−body semi−empirical mode for body−centre cubic transition metals[J]. Philo. Mag., 1987, 50(1):15-30.
[44] A Landa, P W ynblatt, A Girshick. Development of Finnis−Sinclair type potentials for the Pb, Pb−Bi and Pn−Ni system,application to surface segregation[J]. Acta mater., 1998, 46(8): 3027-3032.
[45] C S Barrett, T B Massalski. Structure of Metals (3th) [M]. New York: Pergamon Press, 1980.
[46] C Kittle. Introduction to Solid Physics. 5ed. [M]. New York: John Wiley & Sons Inc, 1976.
[47] D Schumacher, W Schilling, J Diehl. Vacancies and Interstitials in Metals [M]. Amsterdam: North−Holland, 1970.
[48] E A Brandes. Smithells Metal Reference Book [M]. London: Butterworths, 1983.
[49] 刘英. 基于第一性原理计算的原子间相互作用三体势及其应用[D]. 北京：北京科技大学, 2002.
[50] 张邦维, 胡望宇, 舒小林. 嵌入原子方法理论及其在材料科学中的应用−原子尺度材料设计理论[M]. 长沙：湖南大学出版社, 2002.

地址：北京市百万庄大街22号
邮政编码：100037
电话服务
社服务中心：010-88361066
销售一部：010-68326294
销售二部：010-88379649
读者购书热线：010-88379203
网络服务
教材网：http://www.cmpedu.com
机工官网：http://www.cmpbook.com
机工官博：http://weibo.com/cmp1952
封面无防伪标均为盗版

上架指导 工业技术／工程材料

ISBN 978-7-111-39868-4
策划编辑◎张秀恩 ／封面设计◎陈沛

ISBN 978-7-111-39868-4
9 787111 398684 >
定价：39.00元